Philosophy Bear

Live More Lives Than One

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

LIVE MORE LIVES THAN ONE

The best essays of Philosophy Bear AKA de Pony Sum, 2018–2021—revised and updated
This is open beta 0.1.6—I'm looking for comments and criticism of:

-Style
-Grammar
-Font
-Formatting
-Argument
-Article selection
-Anything else you like

If you're not already subscribed—consider subscribing to my free Substack:
https://philosophybear.substack.com/

Table of Contents

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

Preface

Hello and welcome to “Live more lives than one”- a collection of my essays. This book doesn’t need to be read in order. The best way to start is to proceed to the table of contents, find something that interests you, and skip to it.

Sometimes writing my essays feels like casting off my thoughts, just as rats are cast off a sinking ship.

Sometimes I just hope that my gifts intermingle with my weaknesses (intellectual, moral, emotional, aesthetic) to make something interesting by accident, a kind of literary Miller-Urey experiment.

Sometimes I reread my writing and I feel amazed that the son of a chef and a hospital trolley lady managed to write such pretentious twattery—a true pathbreaker for working-class wankers.

And then, after I finish soothing my self-feeling, I tell myself to stop thinking about me and think instead about something that might do some good.

The book is free, but it took a lot of effort to make. If you get anything out of it, I’d ask that you do one of:

1. **Chip in for its advertising— https://www.paypal.com/paypalme/Livemorelivesthanone**
2. **Share it! Facebook, Twitter, email, Whatsapp, Reddit and many more- all great places to send it.**

Dedications: To those who have helped this project along (in alphabetical order): Amanda, Ben, Chris, Dad, Julia, Kieran, Laurence, Michael, Michael, Morgan, Mum, Nina, Riki, Ryan, Scott, Tzvi and Yitzi, as well as all my other friends who played some role in this, and to all others who I owe gratitude- I do not think you will ever know who you are, but maybe that’s okay. Also to Paul Ignacio- a graphic designer who did the cover at a very reasonable cost and, very, very much thanks to the online readers who proofread.

PART 1: YEARNING

Oh death, where is the antidote for thy sting? Or: Prolegomena to a new philosophy of the Common Task

It's about 2012. A friend of mine, about 30 years old, has just died of sepsis. I loved him, and he has been annihilated. I'm standing talking with another friend of mine who was also close to the deceased. A thought occurs to me. "Do you think we'll ever be able to fix it?" "You mean feel better? That will come with time." "No, I mean bring him back from the dead with technology." My friend looked at me in puzzlement and sympathy, thought for a moment, and said "No, I don't think so."

In the past when loved ones had died I had imagined death as a vast granite barrier which my hands could make no mark on. But what if we could find a ram powerful enough that the wall of Hades couldn't prevail against it? The thought seemed stupid, yet the future is long and holds many technological wonders. How could I be so confident there was no hope? A hundred years ago an eccentric, perhaps insane, Russian philosopher named Nikolai Fyodorovich Fyodorov suggested—on the basis of scant to no evidence—that such a thing might be possible.

I want to emphasise that I am not suffering from psychosis, so I do not really hold that the idea I describe here is viable. Yet I can't help but play with it and ponder it. Didn't we get where we are in part through mad dreams? To cheat a little with metaphors, maybe you need a vantage point some distance from what is possible to see the full scope of possibility.

I have a fantasy. I mean this entirely seriously when I say that I think it is *among the greatest fantasies ever conceived*. There is little vanity here because it is not my fantasy alone. What if we could redeem all of history—I really mean all of it. Give every story a happy ending by bringing the dead back to life. Not just slow or stop the advance of death, but reclaim each territory it has seized from us, and so, at least in a sense, correct every injustice there ever was?

My fantasy is a very old fantasy. It is essentially the fantasy of universal salvation. I'm an atheist, but it is typically a religious fantasy. It receives expression in Mahayana Buddhism and scattered forms of Christianity and Islam. I would bet good money that someone in the Jewish tradition has articulated it, but I haven't found a reference yet. I'm sure it can be found in many other places besides. Apparently it's currently a hot topic in Christian theology (or at least the Protestant strand thereof). You can even find a trace of it in the Bible:

“On this mountain, He will swallow up the shroud that enfolds **all** peoples, the sheet that covers **all** nations; He will swallow up death forever. The Lord GOD will wipe away the tears from **every face** and remove the disgrace of His people from the whole earth.”

Emphasis is mine.

Generally speaking, the vision has been a supernaturalist one. In the absence of the supernatural it seems likely that people dissolve at death, with no directions about how to put them back together again retained in some secret archive. At least if the ancient philosopher Epicurus is any guide, this is what naturalists have believed since there were naturalists. There is at least one exception though—one person who thought salvation might be achieved naturalistically. In the late 19th and early 20th centuries, Nikolai Fyodorovich Fyodorov articulated what he called the Common Task.

Nikolai argued that one of the greatest forms of alienation stunting human potential is that of the living from the dead. The division of the living from the dead is greater than any division between nations or caste. While many transhumanists have proposed abolishing death going forward, Nikolai was nearly unique in proposing a *retrospective* abolition of death. Although a Christian himself, he thought, rather boldly, that it might be possible to resurrect everyone who had ever died using science. Without human intervention, salvation would be partial—only for good Christians, or perhaps only for members of the Russian Orthodox church, but a mechanical salvation was possible. Such a salvation would not just restore all humans to life, but make that life eternal through the marvels of science.

If nothing else, what a sweet vision. There’s the obvious, of course: for a hopeless romantic such as myself, Alexander and Hephaestion, Abelard and Heloise, Antinous and Hadrian, Andromache and Hector, Anaïs Nin and Henry Miller, whatever real couple the story of Apollo and Hyacinth was based on—and that’s just couples with names starting with A & H. But far more important than these, nameless peasant 10,405,771,606 whose story you never heard, even though it was far more tragic. The approximately 5,000,000,000 dead of malaria. The roughly half of all children who never made it to adulthood. The lost and broken who lived a long life filled with ceaseless pain. Can you imagine how excited you’d feel if you thought for even a moment that you’d found some way to fix it all? All the jagged sheet of history with misery scrawled on it, folding into something beautiful. It’s a holy thought—I would love to have met Nikolai. Indeed it’s *just* possible that one day I shall.

What are the scientific prospects for this task? Before we get to that, we need to take a detour through philosophical theories of personal identity.

Personal identity

You step into a teleporter, it vaporizes you. A person qualitatively identical to yourself steps out of a machine somewhere else, with "your" "memories", "your" personality, "your" body, etc. etc. There are two philosophical questions here:

A) Have you survived? Is the creature that stepped out at the other end "you"?

B) If you have not survived, is the outcome, from a self-interested perspective, i) as good as surviving, ii) better than ordinary death but worse than surviving or iii) as bad as dying in any other way?

If you think the answer is yes to A, or no to A but i) to B, then you're in luck. The Common Task *might*, from your point of view, resurrect the dead, (or as good as). If you answered otherwise, then the Common Task is unlikely to work, unless we can find some way to actually pluck the dead from the past. If you're interested in these topics, Google "philosophical theories of personal identity".

With that sorted, let's go on to "the science".

1. **The possibility of a trace**

Nikolai himself hoped that as we gained mastery over the physical world, we would be able to, based on some trace left by the dead, reconstruct them in body and mind. It's hard to say much on this, except that if it is possible, it would require—as best as I am aware—as yet unknown physical principles. Whether you think this makes it vanishingly unlikely, or reasonably plausible, is something of a matter of epistemic taste.

As science has advanced, it is true that we have gained access to traces the ancients never would have imagined—DNA and carbon-14 dating, for example. We now could, in principle, reconstruct the bodies of some of the dead through cloning so long as we have their DNA. While this would not fulfill the great task, it is an example of the advance of science uncovering previously undreamt ways of reconstructing that which existed in the past. There is always the possibility of more such discoveries in the future. It is possible we will uncover some, as yet unknown, natural version of the Akashic records, although there seems no particular reason to hope so.

One trace is fleeing Earth at the speed of light—light. Light goes slower through some media than others and can be refracted, so in principle, it might be possible to capture the fleeing light without exceeding the cosmic speed limit *c*. This could then be

used—again only in principle—to reconstruct events on the Earth's surface. In practice, there may not be enough information left, and even if there were enough theoretically, the engineering problem may be intractable even for galaxy-spanning super-intellects.

2. Nearly infinite simulation

Suppose that computing power turns out to be really plentiful. Maybe we can build computers from subatomic parts, for example. Now suppose we enter as constraints everything we know about the past and conduct simulations of the past, weaving endless quadrillion lives and creating numberless people. At the end of each of those lives we take the persons so created and put them in a digital afterlife. Eventually, for every person who has ever lived, one of those people is going to be arbitrarily similar to them. If you think that someone having had an arbitrarily similar life to yourself existing in the future counts as survival, you will have survived—congrats! The process would likely be vastly more accurate for contemporary humans because the endless gigabytes of what is known about us means there are far fewer gaps to fill in with estimation. But while this may give you and your loved ones better odds, it's cold comfort for the long-cold nameless peasant 10,405,771,606, whose best approximation is liable to be **far** looser.

Looming over all of this, of course, is the possibility that we are in a simulation ourselves. Whether that would make the task more or less likely, or whether it might already be underway, will remain open questions. I try not to think about this too much.

3. Time travel

One easy solution, were it possible, would be time travel. Most plausible conceptions of time travel developed in contemporary physics and philosophy of time suggest that it would be impossible to change the past. That would not necessarily foreclose on us going back and grabbing the data

This is a reasonable review in the Scientific American: "According to current physical theory, is it possible for a human being to travel through time?" of the prospects of time travel by an expert targeted at a lay audience. The conclusion seems to be: ¯_(ツ)_/¯ . I asked my old PhD supervisor, a leading philosopher of time, and she gave the same answer.

4. Something we haven't yet imagined

Do you think we're near the end of discovery, or do you think that there are things as yet undreamt of in any philosophy that will one day be dreamt? Almost every human that

has ever lived would be unable to understand options 1 to 3, so who is to say there isn't an option 4, 5 or 6?

Summing up prospects:

To be honest, none of the above methods are especially persuasive to me, at least for all humans that have ever lived. I can conceive of something like option 2 working for humans who lived post the invention of social media, and for the relative bare handful of humans who left substantial documentary traces of themselves prior to this. Dan Simmons imagines something like this being used to recreate the poet Keats in the novel *Hyperion*. With those exceptions though, I'm not convinced the great task is feasible.

But it might be worth considering the pessimistic meta-induction. The pessimistic meta-induction is as follows: Since historically most of our best science has not turned out to be even approximately true, it seems probable that our contemporary best science will turn out to be not even approximately true. Personally I am sceptical of the pessimistic meta-induction in most areas of the philosophy of science, but a related argument which I call the optimistic technological meta-induction seems more plausible. Past attempts to define what would never be possible through technology were very often failures. This is especially true of technologies which we might see as precursors for the great task. Cloning and the creation of life have both been achieved; projects for creating minds and superintelligences and achieving physical immortality are both underway and will surely be completed at some point if we don't wipe ourselves out. Most of those who have ever lived would not have thought these things possible for mortals.

Technology so often surprises us, and that which we thought impossible happens so regularly, because, in the words of the IRA to Margaret Thatcher: "You have to be lucky every time, we only have to be lucky once". Out of all the harebrained schemes by humanity and its successor superintelligences, only one has to work for the great task—or any task—to be achieved. This is what makes betting against the possibility of anything—especially before we even know the fundamental laws of physics—so dangerous.

If you put a gun to my head and asked me to give you a credence I'd say there's only a small chance of this task being possible, but enough that I wouldn't bet with any great confidence against it.

Practical implications of the common task

There is nothing we can do to make the very unlikely resurrection of the dead more likely. Either it's possible at some unimaginable reach of technology or it isn't. About the only contribution we can make is fighting against humanity's extinction, and we should be doing that anyway.

There are lesser tasks which partially fulfil the great task but are time-sensitive and which we can make a contribution to. For example, ending involuntary death. If anyone takes this essay as a reason to aim at these less urgently, I will personally hunt you down and tell you off.

Maybe I'm setting myself up for heartache in the future, but I like to sometimes use the common task as an organising myth in my life. A sense of what would be the ideal outcome of everything, to measure and assess more feasible alternatives against. The role an ethical-aesthetic organising principle is difficult to explain, but it seems to help. Sometimes, when I'm at my bleakest, it becomes a reminder that no one has ever scientifically proven that everything won't turn out all right.

For a philosophy of the common task

Someone (maybe you!) could write a very good book considering the common task from a contemporary philosophical perspective. The common task raises questions in the philosophy of personal identity, time, physics, ethics and metaethics, religion and many more.

Indeed, you could teach a very good introductory philosophy course using the philosophy of the common task as a springboard. Time travel (theories of time), personal identity and persistence over time, the simulation argument (epistemology & scepticism), duties to the dead (utilitarianism against alternatives), the sceptical meta-induction (scientific realism and its rivals) etc.

Perhaps the most unique question is about the value of resurrection. Let us suppose that we can raise all ~100 billion people who have ever lived from the dead, but that we could instead use those resources to create, say, 200 billion new, joyous lives. Should we do the former or the latter? Do we have a duty to the dead to restore them to life if we can? Do we have a duty to the living not to leave them alienated from the dead? Do we have duties to the dead in a way that we don't have duties to the hypothetical persons we could bring into being?

And if we do start raising the dead, do we raise all of them? Presumably if we had the technology to do this, we could keep everyone safe from everyone else, but morally, does the world need Mussolini or Bundy back? Should they at least serve some kind of sentence before joining everyone else? These problems might seem absurd (they are, really), but it's an interesting way to grapple with questions about the reason and purpose of punishment. The seeming absurdity of punishing anyone under these conditions is one of the reasons I believe that punishment can only be instrumentally good.

Aesthetics of the common task

There is an endless range of poems, t-shirts, mini-series, sculptures, novels, paintings, desktop backgrounds and radio plays waiting to be written on the common-task—or on the idea of a secular resurrection of the dead. At present I know of only two works, neither of which I have read. One of which, which I have forgotten the name of, tackles Nikolai's work explicitly, whereas the other is Riverworld.

Aesthetically the idea is almost megalomaniacal and difficult to pull off. It undercuts a central concept of much contemporary literature and art—death as an inescapable existential problem. Part of secular maturity is accepting the permanence of death, so the idea represents, in some sense, a return to a philosophical childhood. I feel it myself, even as I write this with unusual giddiness.

In the past I've suggested that post-scarcity worlds are very hard to write about because they remove many (though not all) of the obstacles that fuel narrative and that this leads authors to a kind of reflexive anti-utopianism, equating that which is bad for story with that which is bad simpliciter. A post-scarcity, post-resurrection world redoubles these problems. Add to this the human tendency to rationalise even involuntary death as a good thing to cope, and I can only suspect many authors and artists would instinctively oppose the great task. It will always find purchase among some, though. Some human problems do remain in a context without death or material scarcity, and while it would be very difficult to write a narrative about a context like that, I think it could be worthwhile.

Ultimately we have to grapple with religious aesthetics if we're going to try to represent the common task. In a personal capacity, I find myself wondering if I am not trying to reconstruct the Christianity of my youth in a thinly secular context. While I don't believe, I can't help but dream of a glorious τετέλεσται, even if it makes me a sucker.

Fuck death.

Bibliographic note for "Oh Death, Where is the Antidote to Thy Sting": *I had (inexcusably) forgotten an email I'd received from Damian Tatum that mentioned computer simulation as a strategy for resurrection. Although I had forgotten the email because it came during a busy period, I can only assume it influenced my thinking on the matter since the parallel between what he and I describe is strong. Alexi Turchin has also written an essay which covers very similar ground, though in a different way. You can find his essay here: https://www.academia.edu/36998733/Classification_of_the_approaches_to_the_technological_resurrection*

Perspectival fever: On being shot through with philosophical desire

I

I have a recurrent experience where something I've done, or something I am, appears to me to be better than I know it to really be. I pause and try to break through, to see it from the outside, but I can't.

When I write, what I've written often seems to me to be really good. Only I know it isn't, at least not usually, because the kind of "really good" I'm aiming for means moving people and changing minds. I have made over 170 posts, and done plenty of self-promotion. If what I was doing had the power to move a lot of people, I'd have a lot more than 100 daily readers by now. Tempting as it would be to write this off as market failure in the bazaar of ideas, I don't think it is.

I finally worked out the trick, though—the source of the illusion that makes it hard for me to see my writing as it really is. It reads so good to me because what I've written vibrates in harmony with what is in me at the time I wrote it. The insights feel fresh and powerful because they've just impressed themselves upon me. The metaphors seem choice because, almost by definition, I must have been in the mood for that metaphor

when I cooked it. The only time I can see what my writing looks like to an outside observer is weeks after I've published it, when I reread it. In that moment I see it as sometimes pale, sometimes gaudy, sometimes obscure, sometimes basic, but never quite singing the harmonies I recall. This is not because the notes have changed but because my mind no longer harmonizes with them.

I often feel that if I could just grasp what I'd written from a wholly different perspective, from many different perspectives, I'd be a long way to being a better writer. Don't misunderstand me, I have the ordinary capacities for empathy; what I crave is extraordinary capacities for empathy.

Could I reverse the process? Take a mediocre (though not bad) piece of text and think myself into the state the author was in when they wrote it, making my thoughts ring in harmony with it till it reads like a model of brilliance? I'm very interested to try. If you've written something you think is mediocre but which at the time felt brilliant, email me. I'll see if I can't bridge my way to where you were standing.

II

Writing is not the only time I have this experience.

When I was overweight and I looked in the mirror, my stomach seemed big, but my face never seemed pudgy. However, I knew it was pudgy from the testimony of others. But stranger still than that contradiction, I noticed that when I took pictures, my face seemed pudgier in the still image than in the mirror. I took thousands of selfies of myself over a decade, not out of narcissism, but because the discrepancy between the two was maddening. I wanted to understand by seeing myself as a stranger, to become an object for analysis.

It's not always a matter of my intuitive reaction being more positive than what I know to be warranted. When I suffer very bad fits of OCD—the kind of fits that make some psychologists suspect the disease is related to schizophrenia—I find myself maddened by a paradox. On some level I know it's all false, because it's all turned out to be false so many times before, *yet I cannot fully see the world as one who thinks that it's false*. It's not just a matter of feelings (aliefs); the bits don't click together right at the level of belief either. I try to see myself from the outside, but I can't escape my skin. As it once occurred to me in a daydream, there is a storm, there is a boat, I am the storm, I am the boat, and there is no way I can sail free.

I call this clash between what I know and feel, combined with a maddening desire to harmonise, *perspectival fever*. Reading what I've written, thinking through what I fear, and looking in the mirror—this is when it hits. I know something to be true, but I can't see it that way.

III

A woman, who I will not name because she has already endured more than enough public shaming, wore blackface to a Halloween party. This was a confused attempt to parody Megyn Kelly's denial that there is anything racist about wearing blackface to a Halloween party. While her gesture was, by all accounts, intended to be anti-racist, it missed. As far as anyone can tell, she had no idea how poorly her choice of costume would be received until she turned up, at which point it was too late. Two years later (???), The Washington Post ran a story about her transgression, apparently in an effort to cover their ass about something.

I used to wonder if maybe most serious transgression and crime in our world was like this woman who somehow didn't realise she was going to get canned for blackface. People just sort of forget that a certain course of action is monstrous and illegal until it's

too late. “I’m a murderer/rapist/mobster? Huh, well, I never thought of it like that, but now that you put it that way…”

Call the mental state of being unaware that you are about to do something transgressive when it should be obvious *moral blindness*. Anxiety about the possibility that we have suffered, or will, suffer moral blindness is pretty common—among the highly religious, in the anxious, shifting enclaves of this age and in various mental illnesses. So many people are afraid of fucking up and not realising till it’s too late, but people don’t often talk about it, in case they end up looking like a weirdo. In other words, people fear temporarily losing the ability to see their actions from the outside and breaking a rule. What is that fear if not another form of perspectival fever?

IV

Perhaps it isn’t surprising that someone with these strange derangements of mine would come to be obsessed with the following thought experiment.

Someone, call her Jiang, fell into a deep sleep, and woke up proclaiming that she had experienced *all of it*. All of human history. All hundred billion lives, or approximately three trillion years of experience. The weeping of Alexander as he saw there was nothing left to conquer. The passion and fear of the suffragette Emily Davison as she fell under that horse. The moment calculus first clicked for Newton. The plight of Pocahontas in England. Above all of these, though, the nameless and numberless of history and their unending days. A day can be a long time; she has experienced over a quadrillion days.

Jiang—and who knows how much of her remains after this experience, but we will assume she retains her identity—establishes her credentials. She then announces that she wishes to address the world. As she mounts the podium for a speech that will surely be watched by more people than any other in history, she opens her mouth and…

Do you think you know what Jiang will say? Not about everything, of course, but maybe you think you know what part of the message is? Stranger still, do you have a hunch that Jiang will contradict some particular belief of yours? If so, how can you possibly justify your belief? Maybe I'm just uniquely thick, but I sometimes suspect I know what she might contradict me about.

Is the question of what Jiang would say even meaningful? The human mind as constituted isn't capable of processing that much data. Perhaps asking what Jiang would think, absent specifying how she would be modified to make it possible for her to process this total experience, is meaningless. Nonetheless I find myself longing, almost painfully, to know what the sum (product?) of all human experiences would be. Perspectival fever on a total scale.

V

Another related question that entertains and torments me. Let's define a "grand convocation," as a hypothetical process in which all the people living in a polity were gathered to decide how a polity would be governed henceforth. Somehow there is allotted for them infinite time to speak and debate, and their capacity for boredom is removed. Each of them can address the whole as much as they like. Do you find yourself fascinated by the unknowable question of what they would decide? What if we vary it—for example, by binding them all to truth in their deliberations, or by greatly enhancing their intelligence.

(Surely you must relish the thought of what they might sweep away?)

Or what about a grand convocation of Jiangs? What if every person in the assembly experienced the life of each other person? All approximately 330 million Americans (or substitute any other state, or substitute the whole of humanity) living the lives of the

approximately 330 million Americans, on top of their own, and then—and only then—hammering out a consensus on how the US should proceed.

VI

Here's another philosophical fantasy for you, this one more exhibitionist than voyeuristic. Have you ever dreamed of presenting your whole self to others? Of giving them a copy of a biography of yourself written by God or a Culture Mind or something? Something that tells the whole story? Or maybe just having the magical ability to make others know you are telling the truth when you are, in fact, telling the truth—so you could tell them everything and answer their questions with complete confidence they would believe you?

Is it a fantasy of forgiveness? Of letting others see you have done wrong, but showing them so much of yourself that they can't help but see it in context? Or is it a fantasy of connection? Of releasing the ache of a loneliness so deep you had forgotten it was there? It's all of these perhaps, but it's also perspectival fever—the desire to be seen through other eyes is, in part, the desire to see yourself through other eyes.

"The Ballad of Reading Gaol" as a rejection of all law and politics

For he who lives more lives than one

More deaths than one must die.

-The Ballad of Reading Gaol, Section III

They think a murderer's heart would taint

Each simple seed they sow.

It is not true! God's kindly earth

Is kindlier than men know,

And the red rose would but blow more red,

The white rose whiter blow.

-The Ballad of Reading Gaol, Section IV

1.

When I was young a number of horrific experiences convinced me that I could either choose to be wholly on the side of humanity—all of humanity—or a misanthrope. I chose the first option, although I fall short constantly. Trying to explain how that commitment to being on the side of humanity works on the level of feeling—to show how certain ideas are emotionally and aesthetically coherent with each other in order to create a harmony in how I feel about humans in general, is what led me to write this essay.

We'll get to "The Ballad of Reading Gaol" soon, as promised in the title, but before we do I want to take a detour through the Gospel of John. (Don't worry—I'm an agnostic and this isn't going to turn into a religious essay.)

One of the most famous passages in the New Testament is the story of the woman taken in adultery. You may remember it as the story with the line: "Let he who is without sin cast the first stone". It's in the Gospel of John:

"[…] Jesus went to the Mount of Olives. Early in the morning he came again to the temple. All the people came to him and he sat down and began to teach them. The scribes and the Pharisees brought a woman who had been caught in adultery; and making her stand before all of them, they said to him, 'Teacher, this woman was caught in the very act of committing adultery. Now in the law Moses commanded us to stone such women. Now what do you say?' They said this to test him, so that they might have some charge to bring against him. Jesus bent down and wrote with his finger on the

ground. When they kept on questioning him, he straightened up and said to them, 'Let anyone among you who is without sin be the first to throw a stone at her.' And once again he bent down and wrote on the ground. When they heard it, they went away, one by one, beginning with the elders; and Jesus was left alone with the woman standing before him. Jesus straightened up and said to her, 'Woman, where are they? Has no one condemned you?' She said, 'No one, sir.' And Jesus said, 'Neither do I condemn you. Go your way, and from now on do not sin again'"""

Now my friend Karl Hand, biblical scholar extraordinaire, assures me of two things. Firstly, there is almost no doubt that this passage is a later addition, written by another author. Secondly, among the relatively small number of scholars who defend the authenticity of this passage, most are conservatives. However, in my research, I found that, while evangelical and fundamentalist Christians generally defend the whole of the bible, on the grounds that God would not let his word be polluted with error, there is a small grouping of far-right cranks who argue that this passage is, unlike the rest of the Bible, inauthentic. The, uh, always interesting source Conservapedia has it:

"Historians and scholars agree that the story of Jesus and the woman caught in adultery is not authentic and was added decades later to the Gospel of John by scribes. The story was almost certainly added for the purpose of Democrat ideology: if no one who has sinned should cast the first stone, then the message is that no one should punish or even criticize sinners. It is also clear from the writing style that this story was added later."

It is most curious, surely, that the very same people who have defended the literal accuracy of the Bible, even to the extent of claiming the world is 6000 years old, are suddenly astute textual critics when it comes to this passage? How overwhelmingly threatening it must be, to be the sole portion distressing enough to move these arch-conservatives away from the doctrine of biblical inerrancy.

2.

The reason why the strange conservatives at Conservapedia are keen to disavow this, and only this passage is that it proposes, more or less explicitly, that because we all share in the same sinful nature, none of us has the right to punish another. Such a perspective, however impractical it may be, is a conceptual threat to all systems of authority, laws, hierarchy, and ultimately even to organised society. Nonetheless, I think it's one of the best wishes anyone has ever made.

The two oldest functions of government are criminal punishment and defense of territory. This last category might even be seen as a special case of punishment—deterrence through the use of incentives. It's often said that the state is defined by a monopoly on violence; well, the most fundamental form of that violence for the state is punishment. This story of a woman, her accusers, and God become flesh cuts against the very heart of government, conventional morality and capitalism. It is, in the purest, most glorious, and sadly most impractical sense, anarchist.

The same radical message appears in many places, but few as eloquent as "The Ballad of Reading Gaol" almost two thousand years later—but here it comes with a twist.

3.

In 1895, Oscar Wilde was sentenced to two years of hard labor for "gross indecency with another man". He spent much of his sentence in Reading Gaol.

While at Reading Gaol he watched, appalled, as Charles Thomas Wooldridge was executed for the crime of slitting his wife's throat. Oscar Wilde was a humanitarian, an anarchist, a socialist, and a man who never softened to the world's cruelties. The idea of executing anyone was truly indecent to him, and he saw the hypocrisy of a violent society punishing violence.

After being released from prison he wrote "The Ballad of Reading Gaol."

4.

"The Ballad of Reading Gaol" is a poem, and therefore its content cannot be distilled into a list of "points". As Harold Bloom once said, the meaning of a poem could only be another poem. Yet there are clear themes which, however superficial it may be to do so, we can grab and isolate.

Where "The Ballad of Reading Gaol" differs from the story in John of the woman taken in adultery is that it proposes two reasons why punishment is fundamentally indecent. These reasons are in tension with each other, but not, I think, ultimately contradictory. The first reason is that we are all fundamentally sinful in nature, so whoever performs the punishment is implicitly claiming to be fundamentally different from the punished in a way which just isn't true. This reasoning can be found in the story of the woman taken in adultery.

The second reason it gives isn't so obviously present in that biblical story. People are noble and beautiful, and whatever their flaws, don't deserve the dehumanisation, agony and humiliation that comes with punishment, at least as it is practiced in our society. Describing the prisoners coming out after the morning of the hanging:

"And down the iron stair we tramped,

Each from his separate Hell.

Out into God's sweet air we went,

But not in wonted way,

For this man's face was white with fear,

And that man's face was grey,

And I never saw sad men who looked

So wistfully at the day.

I never saw sad men who looked

With such a wistful eye

Upon that little tent of blue

We prisoners called the sky,

And at every careless cloud that passed

In happy freedom by.”

Or:

“They hanged him as a beast is hanged:

They did not even toll

A reguiem that might have brought

Rest to his startled soul,

But hurriedly they took him out,

And hid him in a hole.”

The passages where he describes the mourning of the prisoners for Woolridge before and after he dies are beautiful. The contrast between the men’s shabby surrounds and the glory of their souls as they keep a vigil on Woolridge’s behalf rends us:

“The Warders with their shoes of felt

Crept by each padlocked door,

And peeped and saw, with eyes of awe,

Gray figures on the floor,

And wondered why men knelt to pray

Who never prayed before.

All through the night we knelt and prayed,

Mad mourners of a corse!”

5.

It is possible to believe in both bits of reasoning. People are too beautiful and important to be brutalized, and too fallen to administer punishment without being hypocrites. They’re not logically inconsistent, and I don’t think they’re aesthetically or emotionally inconsistent either. Just like a sufficiently skilled art work can contain moments of appalling ugliness alongside tremendous beauty without those “cancelling out”, so too are people woven through with glory and horror. Too beautiful to be judged, and too ugly to judge something as glorious as a human.

6.

I am speculating here, but I wonder if there isn’t something erotic or romantic in Wilde’s outlook on Woolridge:

"And I knew that he was standing up

In the black dock's dreadful pen,

And that never would I see his face

In God's sweet world again.

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

Like two doomed ships that pass in storm

We had crossed each other's way:

But we made no sign, we said no word,

We had no word to say;

For we did not meet in the holy night,

But in the shameful day."

Now, this is further stepping into the realm of pure speculation, but I wonder if that romantically charged perspective on Woolridge wasn't a path by which Wilde humanized him—saw past the horrific thing he'd done? Romantic and erotic energies have this power—to randomly connect us with, and make us sympathizers for, people we would otherwise despise, or at least try not to think about . This is a side of the erotic we don't often consider. We often conceive of eroticism as turning people into objects in our mind, but what about its capacity to make us sympathizers? Sometimes this power takes on a sinister or at least ambivalent aspect—like the people who fantasize about serial killers and court them in prison. Sometimes it is exalted in literature, as in *Romeo and Juliet:* Would a rose by any other name not—etc. etc.

7.

Obviously, a conservative will find much to disagree with in the poem, but “The Ballad of Reading Gaol” is an uncomfortable read, whatever your political orientation. I’m all for mercy, but as someone who thinks women have historically had a rough deal, I wasn’t comfortable with Wilde’s seemingly blithe dismissal of Woolridge’s murder of his wife:

“Yet each man kills the thing he loves

By each let this be heard,

Some do it with a bitter look,

Some with a flattering word,

The coward does it with a kiss,

The brave man with a sword!"

"Well sure, he brutally murdered his wife, but in a funny sort of way, doesn't every man kill his wife?" To which the answer is no. There is a very important sense in which the vast majority of men don't kill their wives—the literal sense. Is Wilde playing with words here to minimize a gross act of violence against a woman?

Perhaps. But there's also a sense in which Wilde's sentiment can be read not as minimization of what Woolridge did, but maximization of the emotional violence inherent in a certain sort of marriage. In this regard, this stanza might be read not as an apologia for Woolridge, but as a biting critique of patriarchal marriage. I'm not fully comfortable with this defense of Wilde, but we shouldn't feel comfortable about art.

To fully draw out the critical power of the poem, we must remember that there are four victims in it. The first is Woolridge; the second is Wilde and the prisoners collectively; the third is the collective warders, doctors, and reverends of the prison who are brutalized by what they do; and the fourth, and most gravely wronged of all, is Laura Ellen/Nell Woolridge, murdered by Thomas Woolridge. Having recognized the victims, we then need to consider the possibility that simply because they are human, not a single one of them deserved what happened to them.

The passage also has to be read alongside Oscar Wilde's own life. Wilde was aware of the sour face of love. Love sent him to prison and ruined his health and his reputation.

8.

Because I like to make up words, let's call generalized opposition to punishment antipoenaism, from "poena" which is Latin for "punishment" and "anti" which is Latin for "anti". Could antipoenaism ever be viable? Is antipoenaism the sort of idea which depends for its interest on whether it is, or ever will be, viable?

No. Antipoenaism is pretty obviously not viable with the world the way it is—some people need incentives not to do bad things. However, it could be viable in a future where we have the technological capacity to restrain the violent without removing their liberty (see Iain Banks' concept of the slap drone) or to cure the violent of their violent tendencies.

But I think antipoenaism is an idea that holds power even in a world where it is not feasible, and should hold that power to shock and shame us all. Jesus' provocation, "Let he who is without sin cast the first stone", holds our imaginations even now. We need a compass that points us towards utopia, even if we can't make it there, and even if it can't be real—you won't see the world as it is without crazy dreams of what it could be.

9.

It's very interesting that the greatest piece of work by Wilde is "The Ballad of Reading Gaol." Wilde was an aesthete, holding that art should be for its own sake—the sake of beauty—and not to serve pedagogic, political or moral purposes. How weird then that his best and most passionate work brims with moral significance and feeling. The chronic ironist driven by circumstances to express real passion is a potent thing (happens all the time on Twitter). I wonder—and this is pure speculation—if Wilde's aesthete sensibilities weren't like a shell to contain his powerful moral sense, which perhaps he feared might be, in today's language, "cringe". When the physical, emotional, and moral torture he had experienced finally burst through that qlippah, his best work emerged.

10.

I don't know if Wilde was, in any overall sense, a good person, I haven't studied his life closely, and even if I had, I am no judge of souls. But it is unbearable to think of what happened to the spirit, at once both kind and soaring, present in this poem. Fuck you to those that valorize the sort of society that did that to Wilde.

On Klutzes

On *Chapo* the other day, Amber A'Lee Frost raised an interesting point. Regarding the novels of John Steinbeck, she suggested that he was feeding into a national conversation happening at the time he was writing. A conversation about what to do about people who just aren't very good at things, but in a way not generally recognised by the welfare state. We will call these people klutzes. Klutzes are not simply unlucky in the labour market; rather, their talents make them poorly suited for it. Klutzes may or may not have disabilities, but their disabilities are not of the sort that can be used to make a case for social assistance under the current rules.

A personal essay tack is required here. Klutzes are a topic of great interest to me because I consider myself a klutz. I'm dyspraxic physically—I have poor fine motor coordination. Cognitively, I suffer from an executive function disorder that makes me chronically absentminded, and while my spatial reasoning can work, it works very slowly. These factors led to me being fired from my first ever job as a kitchen hand by my own father (he's a wonderful dad, I was just a really, really bad dishwasher). Compounding these weaknesses, I have periodic bouts of severe OCD that can leave me suicidally depressed for up to a month at random. I have a certain degree of charisma and customer service skills, but nothing so truly exceptional as to overcome these handicaps. My skills could be summed as:

1.A decent but unexceptional work ethic

2. Reasonably well mannered and understanding, though with a somewhat off-putting tendency towards eccentricity

3. A capacity for research (though with a tendency to distractibility)

4. Writing skills

5. The ability to walk very long distances

There are certain jobs which fit this profile of skills, but unfortunately they're extremely competitive. Consider, for example, being a writer. I have made pitches to many institutions and publications, including the People's Policy Project, Current Affairs, The Guardian, New Matilda and The Conversation, and not gotten a reply. At the risk of sounding bitter (and I am) I've been told that a lot of media organisations aren't interested in cold pitches. If you haven't already got a portfolio of published work, you need to know someone—or so the story goes. The dire odds of me making a living in media are further lengthened because: I'm eleven shades too left for the centrist establishment which holds 99% of the money in media and—this essay excepted—I'm generally averse to personalising my writing or marketing aspects of my life and identity. No doubt I could overcome these barriers were I a genius. I'm not a genius, and I shouldn't have to be.

I'm a failson, but not the son of anyone important.

I have considered a lot of options. Onlyfans? I'm not ugly, but I'm a little too chubby—and even if I lost the weight I'm nothing special. Patreon? I don't have a large enough audience yet (I doubt I ever will). Starting a business? Most of them fail, and with my absentmindedness I'm more likely than most to fuck it up. Maybe you're thinking that if you had my skill set you could do very well for yourself—that I simply

lack a certain can-do attitude. Perhaps you're right. If so, consider "positive mindset" as just another thing I'm a klutz at.

The result is that, while completing a PhD, I work a job that I am unsuited to. I hang onto it by the fingernails—through people skills and a contract that makes me very hard to fire. Once I finish the PhD and my scholarship money dries up, I'll make a -probably doomed—attempt to get an academic job. When that fails I don't know what I'll do next.

The klutz is mostly invisible in our culture. The few exceptions are klutzes who manage to find some special niche for themselves. The absentminded professor, for example—a trope that some people have told me I fit. Really, such quasi-klutzes are the lucky ones. The real down in the dirt klutzes, the ones that have it very hard, don't have any marketable exceptional strengths to offset their weaknesses. They turn up in our culture from time to time. For example, I don't know a lot about incels, but if my friend and host of the podcast *The Conditional Release Program,*, Joel Hill, is right, many incels are klutzes with the additional disadvantage of being conventionally unattractive and hating women.

None of this is to say that klutzes lack talents altogether; they just lack the marketable ones.

As I mentioned, I am one of the lucky ones because my klutziness is incomplete. I'm just organised enough to fake it and hold down a job. While my parents are not rich, they are very generous and supportive. I suspect I would have spent at least a little time homeless and couchsurfing with friends if my parents weren't supportive. If I were charmless or just a little bit more disorganized still, my parents would have to stretch their generosity further.

I believe there's a lot of us klutzes around, though I can't prove it. There is certainly an outsized number of partial klutzes like me in academia.

Probably, since a young age, many klutzes have been told that they are lazy, because a lot of the time that's what being a klutz looks like from the outside—e.g. if you have an average or even somewhat above-average work ethic and are a klutz, you will look lazy to others, because you get through work more slowly. "He couldn't possibly have forgotten again, he must just be too lazy to do it." "He couldn't possibly have broken something again just by clumsiness—he must not be paying attention".

What I'm trying to do here is articulate a new group—to impose a new condition of being into your consciousness. The klutz. I'm not going to propose specific solutions here because that isn't the point—it is not beyond the wit of *sapiens* to provide a decent life and meaningful opportunities for contribution to klutzes. It's a matter of will. Instead I am doing what I most hate: *awareness raising*.

So, conditional on us not forgetting the time or location, let us rally. Upon our badly homemade banners, let us raise the sign *omni homini habeat valorem*, which, if my amateurish Latin translation is correct, means "every human has value" (and if it is not correct, is that not much the better?). Let us fight for a dignity not conditional on technique or power. Arise, my broke and broken siblings! Let us chant, "useless, not worthless".

Existential tragedies—a partial list of the fundamental complaints of being a person.

We all know there is a suite of "existential" tragedies inherent to human existence. Defining the concept precisely is impossible, but I would say roughly that an existential tragedy is a tragedy that arises from very basic and universal, or near-universal, aspects of our experience. I thought it could be interesting to list them. Let me know if you can think of any others:

-The inevitability of one's own death

-The inevitability of the death of those one knows and loves

-The probable unknowability of many important cosmic truths

-The apparent existence of meaningless suffering serving no higher purpose

-The inherent trade-offs around what can fit into a single life

-The possibility of unrequited love (and I certainly don't just mean romantic love—unrequited familial love is usually worse)

-Our total lack of control over the most important factor of our lives—the circumstances of our birth

-The unequal and random distribution of talent

-Our inability to consistently embody even our own idea of the good

-The erosion of treasured (or simply important) memories by time

-The erosion of vitality and beauty by time

-The unknowability of the full results of our actions

-The privacy of experience even when we wish it were otherwise

-The frustration of words that can never fully convey what we mean

-The perpetual possibility of being disbelieved, even when we are speaking important truths about ourselves and our lives

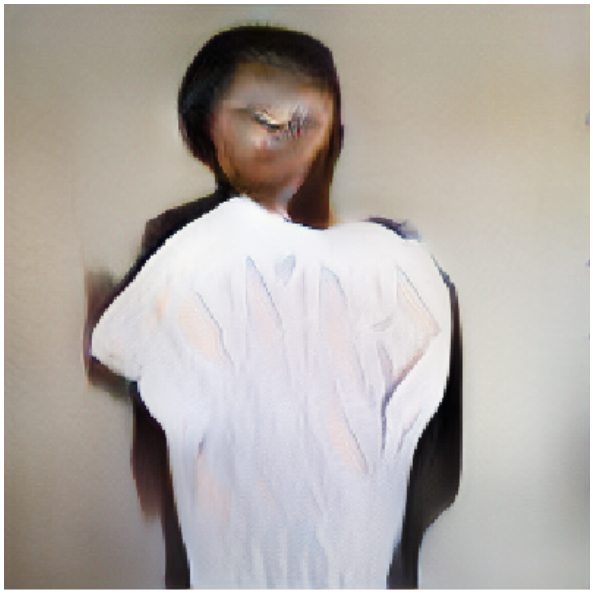
You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

-The possibility of being wracked by want for something which is impossible—not merely improbable but totally impossible.

-Our actions are irrevocable—in a few lucky cases we might be able to prevent or fix all the harm we caused, but even then, this does not undo the action itself

Artificial intelligence dreams images to accompany Sufjan Stevens lyrics

Continuing this blog's project of creating Sufjan-themed art using AI, I've automatically generated a collection of images to accompany Sufjan lyrics. Using this Computer Vision Explorer —Text-to-Image Generation, I gave a computer program many Sufjan Steven lyrics and picked the best results.



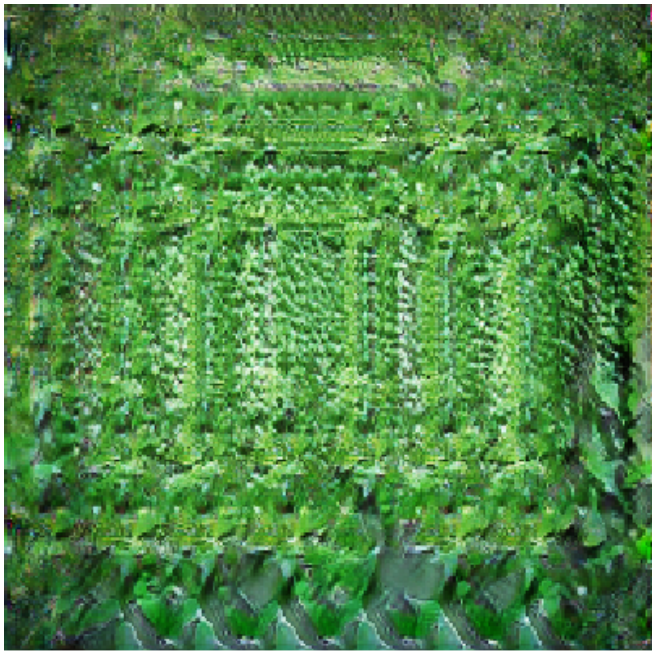
I was dressed in white

You stare at the sun to see the sublime

The lion and the lamb were restored

His father was a drinker

In the tower, above the earth



All things grow

I'll find sleep, I'll find peace

It's your own damn head on that plate

Justice delivers its death

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

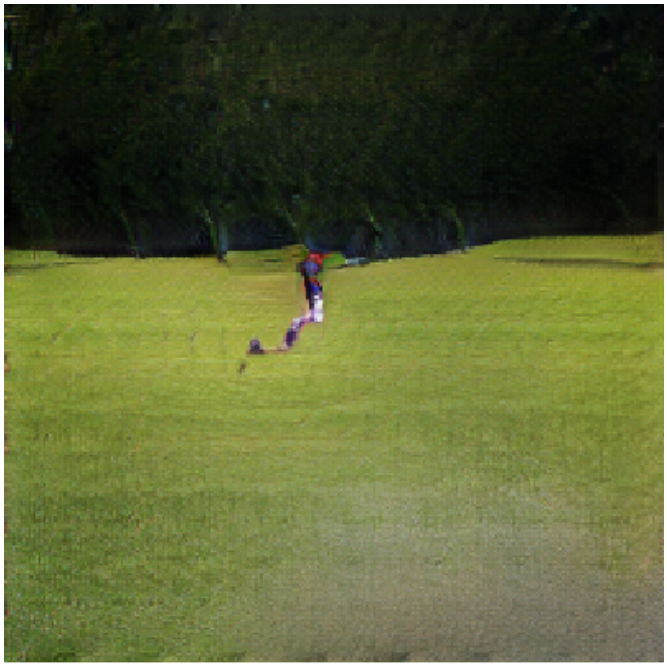
Blackbird on my shoulder

I built your walls around me

All of me wants all of you

Signs and wonders

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear



My friend is gone, he ran away

We were ashamed of her

We're all going to die

The Culture novels and the deaestheticization of politics

I.

You know that old cliché in science fiction? The hardy explorers find a society that appears to be a utopia—but not all is as it seems!

In many ways, the Culture novels by Iain Banks are an inversion of that. Characters either live in, or encounter, the Culture—a society that appears to be paradise and really is. However, the protagonists rebel against it—seeking a world where maybe things are a bit tougher, but damn it, there is glory! They convince themselves that the Culture is a fake utopia. Tragedy results.

Characters embodying this trope include Horza from *Consider Phlebas;* Gurgeh from *The Player of Games* (long fantasises about joining Contact to escape ennui, although he has to be pushed in the end); and Zakalwe in *The Use of Weapons*.

It extends throughout the Culture series, but I think is especially pronounced in the first few books. A case could be made for many other characters—for example, Ziller in *Look to Windward,* although he is content mostly to complain.

Some of them oppose the Culture altogether, like Horza; others like Gurgeh want to go somewhere else for a little while, imagining great adventure. Some of them die; others escape relatively lightly with high to moderate trauma. All suffer because they want existence to have more friction than the Culture offers. Implicitly or otherwise, they romanticise suffering, only to realise in the throes of suffering that there is little romantic about it.

"No, Andrea: Unhappy is the land that needs a hero."

-Bertolt Brecht, *Galileo*

II.

It is difficult to tell a story in a society in which there are no objective causes of suffering. No risk of violence, no ill health and no risk of poverty, not even a risk of unhappiness unless you are unwilling to use the appropriate chemical correctives. Even boredom is banished for all but the most jaded tastes. There are a few possible stories that remain: comic sitcoms or romantic comedies, tales of unrequited love or a memoir of the grief of those left behind when someone voluntarily dies. These are real possibilities to be sure, but far fewer than the stories one can tell in a world like ours, still bound up in the realm of necessity. Banks, for his part, doesn't even try to tell a "purely Culture" story. Every Culture novel is about the Culture encountering other societies that are not so utopian.

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

I think these difficulties are the reason why so many authors instinctively rebel against utopias, and why there are so many fake utopias in fiction which are unmasked as hidden dystopias. Authors instinctively don't like utopias, because they make bad stories. Thus fiction writers, consciously or otherwise, judge the sociopolitical structure of societies on aesthetic grounds, equating poor soil for narratives with poor soil for human flourishing—in the words of Walter Benjamin, they aestheticise politics.

III.

It's a shame that authors instinctively aestheticise politics, because this reflex, as Benjamin argues, is one of the wellsprings of fascism.

This was true in the time of the original Nazis and it's true of the cheap knockoffs we have now. If you've ever seen Nazi memes, they're an ideological mess, but one of the themes that comes through is an obsession with beauty and good narrative as a political goal. From talking about "the ancestors" (always romanticised beyond recognition), to the content of the fourteen words, to bromides about honor that clearly come from the pages of boys' own adventure books (and of wars imagined), to complaints about "decadence" which basically boil down to "I don't like looking at it".

In this regard I do not mean to suggest that a bunch of science-fiction and fantasy authors are closet fascists; I'm sure they're good centre-leftists and all that. I'm not judging them either—it really is very difficult to write a good story set in utopia. Nonetheless, we must recognise some basic truths. Suffering, want and involuntary death are bad, and if we do not destroy ourselves or permanently prevent our own technological progress, we will one day abolish them. Our sharpest want and most urgent action should be to speed this day. Romanticising suffering is cosmic Stockholm syndrome, boot-licking for the brute forces of the universe. Portraying utopias as really secretly evil is a lazy and overdone trope, but moreover, it values a certain kind of narrative satisfaction above our soaring possibilities.

In the words of Belinda Carlisle, let's make heaven a place on earth.

Try to always be kind because you never know when you're incompetent

We are rarely one step from disaster. Most really awful outcomes require at least two things to go wrong. Often, but not always, those two things are being malicious and stupid at the same time.

Hanlon's razor says that you should never attribute to malice that which you can attribute to incompetence. It's not a bad way to think, but it's misleading if you take it too far, because most disasters are caused by a bit of both. For example, a wildfire department is under-resourced because the state just doesn't care enough, and the resources are misallocated within the department because the state is incompetent. A police officer is too incompetent to tell that the suspect is choking to death, and doesn't care about the person enough to stop simply because they are inflicting severe pain.

The good news is that you can short-circuit the synergy between malice and incompetence and often prevent the most disastrous consequences of your actions by trying to be a nice person. It's much easier to stop yourself from being mean than it is to stop yourself from being stupid—you usually know when you're being mean, but you usually don't know when you're being stupid. Most of the time you also don't have to be exceptionally kind to avoid disaster either, just ordinarily decent.

I worked in the Emergency department once in the graveyard shift, as an administration officer. One of my tasks was to buzz the nurse to let them know that a patient had arrived. Since the staff was so small, they couldn't spare a nurse permanently at the desk during the night.

We had a lot of patients who came in for absurd reasons. Sometimes it caused big problems and slowed down care for those who really needed it. Mysterious aches and pains were common, and often the patient would see the triage nurse, wait several hours, see the doctor, and just get referred on elsewhere, because their problem was non-urgent. I know for a fact that other admin staff—not all, but some—used to make their irritation with these patients plain on arrival.
One night we had a guy come in one day saying his balls hurt. He wasn't in severe agony or anything, they were just aching. Mentally I rolled my eyes. However, I believe in being a nice guy, so externally I certainly did not roll my eyes. Instead I gave him a comforting smile and immediately called the nurse. I expected she would have a quick

look and suggest he come back in the morning, as had happened to a couple of other patients that night.

Instead she immediately assigned him just about the most urgent triage category you can get with your heart still working and your limbs still attached. Afterwards I asked about him, and she explained he had a suspected testicular torsion and that the only safe way to treat it was immediately.

Had I rolled my eyes or questioned his coming in, as other admin staff sometimes did, I might have left the poor man a eunuch. It didn't take a heap of kindness to avoid it (I have no delusions about being a kind person), just a tincture of patience.

The moral of the story is that because being wrong feels exactly like being right, you're almost always better off being nice. Kindness covers a multitude of incompetencies, including incompetence you didn't even realise you had.

300 Arguments: A commentary

I recently read Sarah Manguso's *300 Arguments*—a series of 300 aphorisms by the author. A lot of the aphorisms were extremely good; some were things I'd expect any moderately intelligent high schooler to know, but that is always true of collections of aphorisms. Overall, the brilliance was exceeding and the time required was extremely modest. Buy a copy. What I've gathered here is not a collection of the best aphorisms, but rather a collection of the aphorisms I was compelled to make some sort of comment on.

- **It isn't so much that geniuses look easy, as that they make it look fast.**

I remember a story. A man teaches two pottery classes. The first class he instructs to make pots as quickly as possible. The second class he instructs to make pots slowly, taking great care. By the end, the first class has made far more pots, but they have also made better pots, being that they had more practice.

Another story. I was talking about writing with my supervisor, a famous philosopher of biology. I told him that Bertrand Russell had written two thousand words a day. With some gentleness, but also a little scorn he told me, "You are not Bertrand Russell". I think about both these things often, as poles in conflict.

- **You might as well start by confessing your greatest shame. Anything else would be exposition.**

I did this once, to a handsome fellow at a party I quite wanted to bed before reading this book. At first he told me it wasn't that bad, then he slowly grasped what I had told him. He didn't talk to me much for the rest of the evening. Don't take the advice in aphorisms too literally.

To put it in terms of the author's metaphor, there's a reason why the modern style of cutting as much exposition as possible is an acquired taste. It's not really in line with our natural social style. Maybe the brisk modernist style of writing reflects the alienated sociability of our era. Try to talk like a modernist novel, jump straight to the most significant parts, and you'll scare people.

- **A great photographer insists on writing poems. A brilliant essayist insists on writing novels. A singer with a voice like an angel insists on singing only her own terrible songs. So when people tell me I should write this or that thing I don't want to write, I know what they mean.**

This frames it like it's just stubbornness, but the trouble is it's very hard to know whether you're the photographer or the poet, the essayist or the novelist. For example,

Manguso's conjecture that everyone is trying to politely tell her to write something else is pure paranoia.

- **At faculty meetings I sat with people whose books had sold 2 million copies. Success seemed so close, just within reach. At Subway benches I sat next to people who were gangrenous, dying, but I never thought I'd catch what they had.**

There are more places at the bottom of the pyramid than the top in every area of life. Playing the numbers, then, we're more likely to descend than ascend, whatever our game is. Yet we're all temporarily embarrassed pharaohs when it comes to the pyramid of literature. Who am I, the man (?) who dreams of succeeding in writing, to mock the man who thinks he'll one day be a medium-sized business owner with a yacht? I may think of such a fellow as one of Steinbeck's temporarily embarrassed millionaires, but doesn't he have every reason to scoff that I'm a temporarily embarrassed bestseller?

- **What's worse, offending someone or lying to someone? [...] Tell me which, and I'll tell you your problem.**

Offending them, 100%. I recognise my response is unqualified, but we tend to pretend words and offences hurt a lot less than they do. If you don't lie to yourself about how much you can hurt others with a few words, you'll see that you sometimes have to lie to other people.

- **The trouble with comparing yourself to others is that there are too many others. Using all others as your control group, all your worst**

fears and all your fondest hopes are at once true. You are good, you are bad, you are abnormal, you are just like everyone else.

I think of long, pathetic hours on Wikipedia reading the biographies of people who have done the things I want to do, looking at their ages and trying to decide whether I still had a chance to get anything done. It's also worth remembering that we tend to compare ourselves to single facets of others, some corner of intellect, some smidgen of character. We see that at every single point there are many greater people than us. What this misses is that people are matrices of attributes, not lists, and certainly not a single variable.

- **Some people ditch friends and lovers because it's easier to get new ones then resolve conflicts with the old ones. Particularly if resolving a conflict, requires one to admit error or practice mercy. I am describing an asshole. But what if the asshole thinks he's ditching an asshole?**

I often feel terrified about how mean the world is. Then I feel terrified of that perception, because what kind of person sees assholes everywhere? An asshole. No answer for it but to give up the game of assessment and try to love others. Regardless of the possibility that you might be an asshole and so might they.

- **I've put horses in poems, but I've never ridden one. They just seem like such a good thing to put into literature.**

I've put exchequers in. What the fuck do I know about exchequers? It’s interesting how objects—and not always the objects you’d expect—have the quality of poetry or don’t.

- **Within a gesture of apparent perfection, a mortal heart must beat.**

Apotheosis, the moment of rising, is almost always more captivating than descending from heaven.

- **Biographies should also contain the events which fail to foreshadow.**

Unfortunately, we forget many of them exactly because they don't foreshadow, and so they aren't available for biographers.

- **There truly are two kinds of people: you and everyone else.**

I guess all my life I've been in a struggle to suppress the tendency to see this way. On the whole, I still think that's the beginning of wisdom—denying your own separateness—but maybe I fought so hard to suppress this way of seeing that I forgot there's a grain of truth here.

- **When a student surpasses my expectations, I feel proud and betrayed.**

I remember when @Sufjansimone wrote poetry as good as anything I'd written the first time he put pen to paper at my request. I still remember it vividly for a reason.

- **Sometimes ill-informed choices have good outcomes.**

But crucially, they were still ill-informed! We love judging by results, and will look for any excuse, no matter how foolish.

- **Great talents encourage great incapacities, but maintaining an inability to cook an egg or drive a car won't make you into a genius.**

Guilty!

- **My long romance with efficiency has made me miserly.**

Also guilty!

- **A non-specific wish to change the world isn't about the world, it's about you.**

Once more I'm guilty.

- **Having a romantic type is an expression of grief for an original loss.**

What is this, an arraignment? Stop pinning me down like this.

- **Someone I knew prevented me from getting a job. I fantasized about his death. Years later, he was fired publicly and shamefully. Then he was divorced. Then he developed a disabling illness. With each of his new misfortunes, I'm punished further, with secret guilt, for wishing all of it on him, long ago.**

If that is so, did you really ever wish death on him?

- **Having a worst regret betrays a belief that one misstep caused all your undeserved misfortune.**

Ouch.

- **"Horror is terror that stayed the night" & "After I stopped hoping to outgrow them, my fears were no longer a burden. Hope is what made them a burden."**

I don't want to overemphasise this, and it may not be the best strategy for everyone,but one of the best tricks I ever played against my OCD was domesticating it.

- **Bad art is from no one to no one.**

I don't know whether I agree, but this seems like a good one to think about.

- **I write in defence of the beliefs I fear are least defensible. Everything else feels like homework.**

One of the very nicest and most useful types of conversation you can have is with a person you trust well enough to admit this—about yourself, about your writing or advocacy.

- **Our fifth-grade class assembled cat skeletons [...]**

How come every author I admire remembers their school years so much more vividly than me? Am I a freak for treating the first eighteen years of my life as uninteresting and blank? Am I missing out on something universally human, or just universally writerly?

- **With great and solemn portent my teacher announced she would tell us something that her teacher had told her, and that her teacher's teacher had told him, and so on, back to Yeats: "The thing to remember is that no one ever finds out that you don't know what you're doing [....."]"**

In modern wisdom literature, this is one of the most common sentiments. People repeat it endlessly in various ways. Someone could write an essay on why we so desperately need to hear this and repeat it so frantically.

- **A woman starts a rumour that I slept with a man in another woman's bed. Fifteen years later I look her up on the internet and find three DUI mugshots. In the first she's the pretty redhead I remember from college. Maybe a few cracks in the veneer. But in the last one she's obese, ruined. I still don't forgive her. I pity her, but I won't forgive her just for being pitiful. Hating her is an act of respect.**

Suppose that instead of later becoming an obese drunk driver, her problem was that she was unstable and unwell at the time she started the rumour. Suppose she was confused and embittered by the world in various ways, holding onto sanity by her fingernails. Suppose that she made up her rumours in a spirit of desperation. Under those conditions would it have been more merciful to attribute to her the agency necessary to

be hated, or to withhold that? When is it kinder to forgive and when is it kinder to resent in cases of mental illness with at least partially diminished responsibility?

- **I'd like to meet someone whose passage through life has been continuous. Whose life has happened to an essential self, and not been just a series of lives happening to a series of selves?**

I think this is part of what makes people love everything from those "which Harry Potter character are you" quizzes to conceptualising their own mental illnesses. Trying to squeeze down the river of consciousness till under pressure the water becomes ice. Part of the terror is, I think, that if you are not one thing, you can't be a unique thing—you're just like all the other multifarious, situationally defined, essence-free people.

- **Who seems a harmless fool to those above him, is a malevolence to those beneath.**

Indulgence of this sort of thing feels kind to the superior, because you see the harmless fool—you don't see those beneath. Lenience is to be dispensed only to those who don't hold a great deal of power.

- **Interesting people aren't interested in appearing interesting.**

I think this might be the author having a dig at herself. If so, she's being too harsh.

- **I want to ask the happiest person in the world whether it was worth it, all the sacrifices he made in order to become so happy.**

Contrast John Stuart Mill where he says that the happiest people spend almost no time thinking about happiness, and may not even be aware that they are happy. They certainly will not have made sacrifices for the sake of happiness. The author seems to agree later: "Happiness begins to deteriorate as soon as it is named".

- **Whatever you're feeling, billions already have. Feel for them.**

In the case of some of my more odd fears, this isn't quite true, but certainly millions, probably hundreds of millions over the grand sweep of history. One of the things I realised about my OCD—whatever paths I walk down, however alien-seeming, I am not alone. For every fear there are at minimum hundreds who have shared it, as I found in long hours on OCD forums. Realising that was two fifths of the battle.

- **"There were people I wanted so much before I had them, that the entire experience of having them was grief for my old hunger" & "Achieve a goal and suffer its loss".**

There are many other aphorisms of the author's that get at the same point as this one. She's right. Fundamentally, what we want is not fulfilment. What we want is a very specific kind of longing. Often, when I was depressed, what I wanted more than anything was to badly want something.

- **I don't think the lover ever forgets who started out as the beloved.**

I am assured by many wise people that in more or less every relationship there is a slight flaw: One or the other party loves the other more than they are loved. Sometimes not by

very much, but noticeably. I think that, in a long enough time, this might destroy every relationship, but we're human, so “fortunately” we've only got several decades.

- **I like writing that is unsummarisable, a kernel that cannot be condensed, that must be uttered exactly as it.**

There are two ways to be unsummarisable, one usually good and the other often, but not always, bad. 1. Be extremely compact. 2. Be vague, so you can always claim that a summary didn't quite land. There’s also a third way: Make the journey more important than the destination to what the writing is about. That’s the poetry road.

- **After a friend dies young, the story of her life becomes the exposition to a tragedy. This is the central problem of biography.**

The tendency to see endings as a summary—as if we were essays—does violence to the person. But as Aristotle observed, art must be a unity with every part linked together through shared meaning. So are we to do violence to the person or to the art?

- **Those without taste smugly praise the thrice belaureled. Poor taste is something else.**

I think the reason this seems tasteless is because it seems dishonest and plagiaristic, like they've gotten their opinions from a conversation guide. Nothing to fear about praising Shakespeare so long as it's genuine praise. But how would you know that the love is genuine? Maybe best to lie about it—just in case people think you got your opinion from the conversation guide. Perhaps this is why so much aesthetic conversation is in bad faith to begin with? From the very start of the discourse, you’ve got to deny too much

fondness for overly popular stuff, but that popular stuff is popular for a reason, so chances are you've got to lie about how you feel about it. From there, you just go on to further build on a foundation of lies.

- **In a description of some annoying rich kid roommates, Manguso mentions their "inane preppy Marxism".**

This is an interesting trap. There's a lot of very useful and basic truths in Marxism, and it's easy to get caught on them and substitute that for an intellectual personality. Marxism isn't the only thing which can do this, but it's probably the perfection of the form. Other examples include neoclassical economics, linguistic structuralism, psychoanalysis, etc.

- **You aren't the same person after a good night's sleep as you are after a sleepless night. But which person is you?**

I think most thinking people can eventually grasp, at least intellectually, that they are both people. What's harder to grasp is that other people you meet are also both people. Rarely do people wear masks; they're just different in different situations. The fundamental attribution error and all that makes it hard for us to see that other people don't have essences (at least simple essences) than to see it in ourselves.

- **I'd never have guessed which people I'd still know by now.**

If you're older than about 24, write down a list of your 10 closest friends and acquaintances who aren't family. Check whether this is true for you. Then take the

opportunity to contact all of them that you haven't heard from in at least a week. You've lost others before; don't let yourself lose these ones.

- **Instead of pathologizing every human quirk, we should say "By the grace of this behaviour, this individual has found it possible to continue".**

It's my favourite cliché: "Be kind, for everyone you meet is fighting a hard battle". –Ian MacLaren

Brief Reflections

Having read Manguso, I was tempted to attempt a few aphorisms of my own:

- The difficulty with writing aphorisms is that you've got to find a rare enough insight not to be trivial, but self-evident enough that merely stating it is persuasive.
- Aphorisms are innately the most egotistical form for this reason. *"I don't even have to give you substantial arguments to be worth reading"*. I apologise for that. This feature of aphoristic writing explains Twitter.
- If Twitter is a game, the parameters of that game mean that the vast majority of its players *must* be losing.
- Creativity is not about turning off the filter. It's about turning a simultaneous process into a sequential process. Normally we filter our ideas at the same time we make them. By contrast, in a creative mode of thought, we make a reiterating two-step process of creating, then filtering, creating, then filtering... It's like the tug of war between the gods and demons over the naga serpent which churned the ocean of milk.

- It's very hard to find a method for individually working out whether a new thought is A) an original (or at least rare) insight or B) something everyone else has already figured out except you. This is why people in the business of thinking up original thoughts are all paranoid about being laughed at, egomaniacs, or both.
- The furtiveness of a bad conscience—its exact balance of shame and guilt and other fascinating parameters—is the secret master of the world. Secret because it's self-occluding. Definitionally people don't want to talk about it. This means its real influence is assuredly manyfold times larger than its apparent influence.
- Guilt's a pretty dark emotion. It almost invariably shades into *grief that I have to deal with the emotional consequences of having done wrong* rather than *grief that I have done wrong*. The line between guilt and shame is also worryingly thin, as is the line between guilt and fear of punishment. Often the only thing more reprehensible than feeling guilt is not feeling it.
- One of the disturbing aspects of how the media has chosen to cover #metoo is that it is covered as a sex scandal. This means a degree of the interest is prurient. I remember a case broke about a footballer accused of rape in my country. The tabloid in question chose for its accompanying image the footballer in question in a pair of speedos.
- The best argument for loving everyone is that you can't fully love anyone unless you love everyone. We all contain each other’s flaws and capacities, even if only in fragments.
- It's simple enough to get statistics on what percentages of people affiliate with religion. What I want is statistics on what percentage of people believe in god in the same matter-of-fact way they believe in Belgium. I want to know how this has changed over time. I suspect it's more dramatic.
- It's interesting that it's Astrology—rather than Tarot, which is undergoing a revival, despite the frankly superior aesthetics of Tarot. One suspects that the role of astrology in *assigning identities* is part of the attraction, much like

people love the Myers Briggs for giving out labels. But equally important is its allowing you to label others, in a fragmenting world.

The questions that haunt me at 3 in the morning

I think that everyone has unanswered questions that bug them. Recently though, I've been making a point of jotting them down. I've begun to realize how many of these thoughts are reoccurring, and that these constellations of questions define my mental life almost as much as my beliefs.

You may think that you know the answer to some of these questions—you may well be right. You may think that the answers to some of these questions are obvious. Here I would have to disagree with you. There are times in my life when I would have agreed with that about some of them, but I've became less and less confident of them over time.

The Questions

Psychology

- To what degree are people motivated by sadism in everyday and political life?
- Is self report a reasonable measure for variables like personality and happiness?
- What is the optimal amount of caffeine to ingest? How does this vary by life you are trying to live.
- Sovereign citizens believe both A) That the legal system is run by incredibly evil people entrenched in power B) That if they can just say the right sequences of words, those people will be compelled by the rightness of their arguments to let them off various crimes, civil liabilities etc. Obviously this combination of beliefs is irrational—that's not in dispute—but what about it makes it so compelling? Shouldn't the factors that make you feel the legal system is incredibly evil also make you feel like you're not going to be able to sway them just by making a strong argument? Why is this seemingly contradictory combination of beliefs such a potent attractor?
- Would the Milgram & Stanford Prison experiments replicate if run properly today?
- What the hell is going on with the replication crisis? A lot of the failures to replicate are in really simple experimental designs. How much of it is outright fraud? How much of it is pure chance and the file drawer effect? How much is

ad hoc analysis and statistical fishing? How much is the participants giving the experimenters the "expected result"?
- How can we make trying to replicate experiments an honourable and attractive path given the structural incentives of academia?

Criminology

- What percentage of convicted incarcerated people are innocent?
- What would be an acceptable false positive rate in a just criminal law system? How does this apply to crimes which can be extremely destructive, but which are often by their nature extremely difficult to prove "beyond reasonable doubt".
- Under what conditions -if any at all-, and in what ways -if any at all-, should society informally punish people for whom a criminal conviction is likely impossible?
- Are people who have done horrific things (rape, premeditated murder) generally all round bad people? Or is human character contradictory in such a way as to allow at least a significant minority of such people to be, despite it all, good or at least average or not far below in most other areas of life? This may sound absurd to some readers, but a lot of anecdotal evidence tends to suggest it might be true. What is going on here? To what degree do people have consistent moral character?

Misc human sciences

- To what degree are there political or social ideas which, even now, if someone dreamed them up and went to some modest effort to promulgate them would transform the world? Are such remaining unthought ideas relatively few, and their effects mostly modest? Relatively many, and quite a few with great effect? Some other combination? The real underlying question here—to what extent is theorizing about political and social issues an effective strategy for changing the world?
- What would be the social effects of a working lie detector?
- Similar to the question about a lie detector. A wizard waves a wand. From that point onward, no one can lie. Does society A) collapse B) get much worse on net but continue C) continue with surprisingly little change D) get much better on net E) become a utopia?

- Is it possible to throw down the tyrant irony and return sincerity to her rightful throne?

Politics

- Our period is defined by greater political divergence on the basis of age than ever before seen since we started taking polls on this sort of thing. Preliminary evidence suggests so far that this is a *cohort effect* not an *age effect* auguring that the younger generations will not age out of their opinions as they get older—at least not automatically. Will this trend continue?
- To what extent are people who make comically bad takes about politics acting in bad faith—versus stupid or self deceiving? Especially those with the expertise in politics to know better? Let me give examples from both sides of politics. I recently saw a senator complaining about raising the minimum wage to 15 dollars partly on the grounds that when he was young he was paid only 6 dollars. Adjusting for inflation , this turns out to be 24 dollars in today's money. On the other side of politics, I recently saw a news station try to insinuate that privilege is the reason we are paying for the lawyers of many of the Capitol Hill rioters. In cases like this—awesomely stupid takes by people with adequate education to know better—are people **A)** Lying **B)** Bullshitting in Frankfurter's sense or **C)** engaged in feats of incredible self deception or **D)** apocalyptically stupid This might sound more like a complaint than a question, but I mean it sincerely as a question—what is typically going on here?
- Why aren't democratic politics more effective? There are certain issues on which the average member of the public disagrees greatly with the average politician, and yet there is very little movement. Obviously money in politics has something to do with it, but how does money exert its influence exactly? It's easy to get blase about this replacing incredulity with a faux wise cynicism, but if you stop and think, it's weird that 70% of the public can strongly support something, yet it be considered a fringe position among politicians. How much of it is caused by monopoly power wielded by political parties due to existing voting systems? How much of it is due to rational or irrational voter ignorance? How much of it is caused by voters deceiving themselves about what they believe—with their true sentiments closer to those of politicians? How much of it is due to the action of the media? What forms of action by the media (and other ideological apparatus) count here?
- What on earth happened sometime in the mid seventies? The wage stagnation, the incarceration spike, the union membership decline…

- Is neoliberalism a useful concept for understanding this phase of economic and political life that began sometime in the mid 70's? Was it ever useful and is it still useful?

Philosophy

- In philosophy and several other disciplines it seems like people were having more "big ideas" before about 1970. The period from Frege to the tail end of logical positivism and ordinary language philosophy – or maybe even till 1980—seems to have been very fecund. Is this an illusion of hindsight? Or is it possible that we've exhausted the "low hanging fruit" in certain disciplines, and the remaining brilliant theories and questions are harder to find?
- What the hell are qualia? How do they fit into the world?
- There are many positions which I think are good "in an aspirational sense"—as organizing myths and sources of inspiration—but would not work in practice—at least with the current level of technology. These include communism, prison abolition, and to a certain extent, in some moods, anarchism. Is this meaningful in any way, or is it just a glorified way of saying "I wish things were different"—a vacuous waste of energy?
- Is there a compact definition of what it means for a person's life to go well for them which doesn't imply that we should do something absurd—like tile the universe with people enjoying their best moment over and over again?
- A lot of questions in philosophy *seem* like semantic wrangling. Often undergraduates will claim that this or that topic is just semantic wrangling. Why is it so rare for professional philosophers to openly embrace the view that a lot of philosophy is semantic wrangling? One could then argue that's nothing to be ashamed of because some sorts of semantic wrangling and surveys of the semantic landscape are important.
- Is consequentalism self-effacing? Would we consequentialists be well advised not to hope that consequentalism becomes widely appreciated common-sense because this might have various bad consequences (undermining integrity, inspiring a certain callousness, allowing people to rationalize their preferred course of action)?

Economics

- Most actual businesses don't face rising marginal costs—instead marginal costs are typically either falling or constant. How should neoclassical economics be transformed in light of this?
- What lies at the bottom of the Cambridge Capital Controversy well?
- Why do we keep using economic models even though they're all rubbish?
- Why don't we talk more about the above three questions?

Art

- An argument can be made that poetry was once the most important of the arts, yet now it is effectively dead!??!! Is it because poets became too focused on avant gardeism and not enough on classical technique (esp rhyme & meter)? Is it because singing -when done well—is strictly superior to poetry for most people, and technological advances make high quality and/or catchy singing always available? I tend to favour the second explanation, because the first explanation implies that poets are just ""leaving free money on the ground" by not returning to classical technique. Nevertheless, this is a great artistic and cultural mystery. Even the fact that there is not more discussion of the disappearance of poetry is a mystery in itself.
- Why is so little fantasy and science fiction writing literary in the capital L sense? Why has no one written Love in the Time of Cholera except with fae? Ulysses with familiar spirits? Obviously there's magical realism, but I'm surprised more people haven't tried a blend of literary writing and genre fantasy. Maybe fantasy writers are too smart to fall into that trap?

Sufjan Stevens

- In the song "The Mystery of Love" the narrator says that they are "like Hephaestion who died, Alexander's lover", but it seems they are actually the opposite of Hephaestion, who, after all, died still the beloved of Alexander. In what sense are they "like Hephaestion".
- What does Gideon, the biblical judge, have to do with a breakup?
- Why Ursa Major?
- What sense to make of "The Ascension" the title track of the eponymous album? Who is Cordelia? What is a chariot hallucination?

- How to understand the dialectic of a singer whose songs are very personal and whose life is very private.

AI safety related

- How much smarter than a human would you have to be to trigger a singularity?
- Imagine a person started off in good health in a first world country with an apartment, internet access and $5000. How much smarter than a human would they have to be to take over the world in one year?

Medical

- It seems like we should be able to greatly reduce the length of time it takes to become a medical specialist. Indeed the length of time required smacks of industry protectionism & regulatory capture—especially when one considers that the various colleges self-regulate who can and can't access the relevant titles. Yet these lengths of time seem pretty consistent around the world. Is there anyway we could safely slash these times by a third or even a half (e.g. six years to become a psychiatrist out of high-school)? Is the specialty structure itself, where everyone learns to become a doctor and then specializes, the best way to handle medicine?
- Linked to the above—medicine generally has been more resistant to proletarianisation than any of the other professions (law in particular has been smashed, pharmacy even more so). Could medicine be proletarianised and doctors made just another type of worker? Should it be?

Personal

- How can I live an authentic life, true to what I actually think, without being wildly ineffective and getting people offside? Presumably the answer is a compromise where I try to keep the lion's share of both my effectiveness and authenticity by sacrificing a bit of both, but isn't sacrificing a bit of one's authenticity for strategic reasons *massively inauthentic*?

Chess

- How close are we to the ceiling of performance in chess?
- Consider a being that can see all possible game trees in chess, and in particular knows for each move whether it keeps the current game state, (win for white with best play, win for black with best play, draw with best play) or pushes it into a worse game state for its side. It is playing against a superb human grandmaster. Presumably the starting position of the game is a draw with best play. The being then picks from those moves which don't "slip" the game state into something worse for it. Is there a relatively simple way to describe an excellent strategy for this chess super being to increase the likelihood of its human opponent slipping up and making a move which throws the game into a win state for the chess super being? of such a chess super being, E.g. one option, although I think it doesn't work, is "choose those actions which increase the total number of losing options for my opponent, even if she picks the shortest path to a draw".

Autopsy on a dream

0. You're standing at the bus stop as the bus comes, looking for your bus pass. You search pockets and wallet compartments with frustration. At a particular moment, this frustration flips into something else. Previously you wondered how long it would take you to find it, and how much of a fool you must look flailing around for it, but as of that instant, the nature of your problems transforms. You now know you won't find the pass. Of course, it doesn't stop you from searching for the pass, but your search has a fully different character. We need a word for the moment of this flip, the figure-ground reversal of hope and defeat.

1. Many years ago: I am in professor Paul Griffith's office. I have just told him that I try to write two thousand words a day. "That's too ambitious" he replies. "But Bertrand Russell wrote two thousand words a day", I reply. He looks at me with gentle humor. There is no malice in his words, but there is certainty "You are not Bertrand Russell".

2. In Terry Prattchet's discworld novels, it is remarked that million to one-shots almost always work. In astonishing defiance of reason, kindness and prudence we teach our

hearts the same thing. I'm sitting in my office, newly 33 when, perhaps for the first time, I alieve that I'm not going to make it as a writer. I have long believed that I am not going to make it, I'm not a fool, but mostly I did not *alieve it*. I **knew** the odds were too long, but I **felt** I would succeed anyway. We spoil our own hearts, and then we beat them.

3. Ages. Immanuel Kant, *Critique of Pure Reason*, 46. David Hume, *Treatise on Human Nature*, 28. Sufjan Stevens, *Illnoise*, 30. Leonard Cohen *Hallejuah,* 50. Jesus Christ, The Sermon on the Mount, 32. Karl Marx, *The Communist Manifesto* (yes that is his best work), 30. Have you spent hours on Wikipedia doing this too? Googling late bloomers and the like? It's one of those weirdly specific things that turns out to be a very common experience—like worrying that your dead relatives are watching you masturbate.

4.I look down at a referral for a patient in front of me. She is 57 years old. Her conditions include "urinary & bowel incontinence, schizophrenia, PTSD, agoraphobia, anorexia nervosa" and a dozen more. Her medication list has no less than 22 items and is evenly divided between Tolkien elf names, and what I can only assume are the names of dead-dread gods (why would you call it Cephalexin???) She has come to us for treatment of her osteoarthritis of the spine, accentuated by her weight (approximately 400 pounds). I call her, determined to get her into our clinic at once. I am even convinced that maybe getting her booked in is the reason I was born. I get her answering machine twice. I finally speak to her, she explains we can't schedule the appointment too close to the anniversary of her mother's death, since grief will distract her. My heart is breaking, but it's nothing against the shivers in her voice. The only happy emotion I can find is pride in her for staying alive. I book her in and finish up the call. Then I get back to worrying about the fact that I'll never be an accomplished author, vanitas, vanitas.

5.A few months before this, I am browsing the internet when I see some people talking about my blog in a forum. One of them mentions that it had "saved his life more than once" referring, I think, to my posts about OCD. I am confused and humbled. I am with

friends, and I show them what's on my laptop screen like I am a kid with a puppy and show and tell. I have forgotten, more or less, by evening.

6. I hate myself for the kind of myopia typified in 4 & 5, but I'm far from the first to observe that hating yourself is just another form of vanity. Musing over the fact that hating yourself is just another form of vanity is also another form of vanity, and reflecting on that is another level of vanity again, so on and so forth up a conceited ladder to a pointless heaven of self-absorption. Nothing for it but to stop worrying about how vain you are and try and do some good.

7. Even at 33 I am already noticeably dumber in some respects than I was five years ago. Much worse, I would estimate time goes 40% faster, and that's at minimum.

8. As I write this, I think about how a gifted prose stylist would regard these scratchings, I can't imagine they'd see them well. "Well fuck them'", I resolve, "They're the freak, not me".

9. The thirst to create drives out the thirst to know, and the thirst for fame drives out the thirst to create. Once, before all these when I was so young as to scarcely remember, there must have been the thirst to see.

10. "Unhappy is the land that needs heroes". Well sure, but imagine, as is sometimes posited in fiction, that there was a genuinely final confrontation with evil. Good, represented by some hero or small band of heroes, triumphed. After that life was to be wine and cake, but with no glory. Can you imagine how bitterly so many people living in that world would resent our hero? All the more bitterly, because bitterness is most bitter when it can't show itself. All the more bitterly still again because bitterness is most bitter when all morality and convention says it should instead be gratitude.

11. I think it's interesting to imagine what the life of our hero -the last hero—in the above situation would be like. For example, he might spend the rest of his days walking as if the ground might fall from underneath him, in fearful recognition that many people

must be seeking an opportunity to reverse the reputation they had grudgingly given him. On the other hand, perhaps she would enjoy the pleasure of being a tyrant as well as a hero. She might gloat, internally, that she was the one who had prohibited heroism, while greedily drinking the remaining supply. I hope not, but then again, heroes often aren't good people.

PART 2: LATE SOCIETY

Yvne: The forgotten opposite of envy

Yvne (pronounced “Iv-Knee”) is the opposite of envy. Where envy is unhappiness that someone has done better than you, Yvne is the joy and satisfaction that comes from a sense of having done better than someone else. Yvne isn’t joy in having nice things, or having done great things, Yvne is joy in having *nicer* things and having done *greater* things.

If you’ve never heard this word before it’s because I had to invent it. Its closest pre-existing equivalent in English is probably the German loanword Schadenfreude, but this is not quite the same thing. Yvne does not necessarily require that a calamity befall the other person, or that their condition be miserable, simply that your success be in excess of their own.

It’s something of a mystery why I had to invent the word. After all, Yvne is no less common than envy -the rich indulge in it all the time—and it is no less objectionable than envy either. Indeed, I would argue it is worse. Envy pits your interests against those who are doing better than yourself and gives you a reason to drag them down. Meanwhile Yvne pits your interests against those who are doing worse than yourself—it gives you a selfish reason to prevent those weaker and more vulnerable than yourself improving their station—isn’t that more sinister than envy? At least envy is about punching up, yvne is all about punching down. Yet envy is denounced as one of the seven deadly sins, while people very rarely even talk about the concept captured by the word “Yvne”, despite its dangers and ubiquity.

You may have already guessed the reason I think people talk about envy all the time, but very rarely talk about yvne. The most powerful people in society have much to fear from envy, and so wish to condemn it. Meanwhile, these same powerful people enjoy yvne as a secret wellspring of pleasure. Back when religion held more importance, it was employed by the powerful to condemn envy. Now the task falls to economists. By contrast, the opponents of Yvne have never enjoyed the same level of funding.

Postscript, why yvne cannot be identified with self-satisfaction or smugness:

I’ve had some people argue in response to this piece that yvne can be identified with either self-satisfaction or smugness. There are two main problems with this.

Firstly, I can experience yvne even while totally dissatisfied and dismissive of my own achievements. For example “I’m a total loser, but at least I’m doing better than my cousin.”

Secondly, one can experience smugness without even thinking of others. I might feel smug and self satisfied upon completing a log-cabin in the woods, without even thinking about other people.

There are many concepts that are “almost” yvne, and overlap in many, even most cases (schadenfreude is another one, as is pride in general), but none are quite the same thing.

On critical social-technological points

1. Critical Social-Technological Points

I want to introduce the idea of a critical social-technological point (CSTP). A CSTP is a technological discovery or implementation after which the existing hierarchies and ruling class of a society is locked in, in the sense that removing them from power, or even resisting them in any measurably important way, becomes much more difficult.

Saying that a technology is a CSTP is different from saying it is an inherently authoritarian technology—A CSTP generally only threatens authoritarianism if it is achieved in an already authoritarian society.

Examples of possible CSTP’s include:

1. Genetic engineering. Genetic engineering could be used to lock in the existence of an unequal power structure by genetically writing in obedience to the underclass, (or it could be used to lock in vibrant democracy by amplifying the human tendency to treat power critically).

2. Surveillance technology. While the mere existence of widespread surveillance has not locked in authoritarianism a la 1984, the creation of artificial intelligences capable of monitoring all surveillance channels for disobedience simultaneously might. If such technologies are successfully deployed by existing illiberal intelligence apparatuses, the capacity for resisting power is greatly reduced, and those apparatuses might become impossible to dislodge.

3. The creation of artificial super intelligence. Writers like Nick Bostrom have worried about an artificial super-intelligence totally out of control of humanity- gaining a decisive strategic advantage. This is indeed terrifying, but only slightly less terrifying is the prospect of an artificial superintelligence under the control of the wrong segment of humanity gaining a decisive strategic advantage on behalf of its masters.

4. I have previously suggested that, even in the absence of technological unemployment, automation might represent a critical change in the balance of class power and thus a CSTP, permanently removing much of the power of the working class to resist. The idea is that fewer and fewer workers might be necessary for truly essential infrastructure.

2. Past CSTP-lite phenomena

If the thrust of the above is correct, a lot of potential CSTP's are coming up soon. You might rightly wonder if there has been anything like a CSTP in the past? Obviously not in a *permanent* form because there is, at present, no almost omnipotent ruling class, and probably not as sudden and dramatic as some of the possibilities I outlined above—but something along those lines?

While I am not an antiquarian a few possible candidates include:

1. The invention of agriculture generally.

2. The invention of the ox drawn plow

3. The invention of the chariot

4. The discovery of bronze working

3. What must we do?

The nature of CTSP's is that they block our sight as to what the future looks like past them. Technological advances are inherently hard to predict, as are social confrontations—and the combination forms a cataract in our already cloudy view of the future. Nonetheless, I think that, cumulatively, between the four possible cases I described, we have enough evidence to believe that one or more CTSP's are approaching.

The more power ordinary people have, and the more accountability to which elites are subject, the better our chances are. The creation of a culture of real, conscious democracy in which we jointly and responsibly decide the future—in which bodies like the military and intelligence cannot simply always plead secrecy—in which the spoils of advancing automation is something we all share in, rather than being immiserated by—may be something that must happen soon, if it is to happen at all.

The paranoid style in petit-bourgeois politics

I've found that a lot of Qanon supporters and other rightwing conspiracy theorists are relatively financially successful and/or small business owners. The element of rightwing political affiliation is unsuprising on demographic grounds—these people are older, white, Republicans so why should anyone be surprised that they often own jet-ski dealerships and diners? But what about the conspiratorial element of these worldviews? Is there a way in which the lifestyle of a small business owner might lead to conspiracism?

It makes a lot of sense when you think about it. Put yourself in the position of a 60 year old, white property investor/car dealership owner. you're doing pretty well in the financial aspects of the game of life, but like a lot of people who are doing fairly well (especially older people who are doing well), you feel like you haven't gotten all that you *deserve* dammit!

But you're in a double bind. You can't question the rules of the game wholesale, because you want to believe that the game is in essence fair, insofar as it has put you ahead of many other people. That means any kind of left-wing critique of the justice of the social rules is out. You can't even really imply that the problem is ordinary breaches in the rules, like tax-evasion, because you've probably done those things yourself.

So if you can't question certain rules of the game, one strategy to explain the discrepancy between what you have and what you think you deserve to have is to imply that other people got ahead of you by breaking the rules- and not in the small ways you do. Elites must be totally disregarding the rules. Since normally breaking these rules is punished by society, there has to be some way a set of elites are circumventing them en masse. A secret, en masse violation of the rules sounds a lot like a conspiracy.

Maybe you think they're tied together by a blood that most people don't share (anti-Semitic conspiracies), maybe you think they're tied together by an ideological project to circumvent and eliminate the economic rules altogether (anti-communist

conspiracies), and maybe you think they're tied together by supernatural pretensions or even real supernatural powers (conspiracies about the occult and elite Satanism)

Complicating the story I've told though is the reality that these people are not entirely wrong. The hyper-successful absolutely do breach the rules all the time, often through conspiracies, although this isn't the only, or main reason that these people are more successful that the conspiracy mongering petit-bourgeois.

This all might sound pretty niche, and I guess to a degree it is, but to the extent that both conspiracy mongering and the petit bourgeois are linchpins of contemporary MAGA-dom, understanding their overlap is interesting.

TLDR: The conspiratorial style in petit bourgeois politics arises from a fundamental conviction that they are hard done by—that you haven't got all you deserve. Naturally if you think you've been hard done by, you'll want to understand why. However, in explaining why you have been hard done by, you do not want to admit any fundamental problem in the underlying social order, because you have profited from that social order, so instead you must appeal to the idea that the social order is being circumvented. The most natural way to explain mass, gross, but hidden circumvention of the social order by everyone who is ahead of you and doesn't deserve to be is to posit a conspiracy.

Twitter is a reverse panopticon: The internal agent

Foucault in a frankly over cited discussion refers to Jeremy Bentham's panopticon—a hypothetical prison in which one never knows if one is being watched because there's one way glass everywhere. The possibility of being watched instils behaviour similar to if one actually were being watched all the time. Eventually (and this is somewhat my own gloss) the hypothetical watcher becomes internalised within the "watched" person's own psyche, potentially unreal, but structuring their behaviour and attitudes anyway.

Twitter (and I suspect other mediums like Tik Tok & Youtube) is like this, except instead of *fearing*(1) that someone important might be watching we are *hoping* someone important might be watching.

This hope reshapes all online consciousness, eventually becoming a kind of internal, clout seeking voice. If the superego is an internal disciplinarian that arises from the external discipline of our parents, I call this figure the internal agent and picture him as this guy, the first thing I found googling "sleazy Hollywood agent":

The Rosetta stone to Twitter is that most of the accounts you see on your feed are hoping to be discovered. I held back from saying this for a time because I was afraid that it was really only a reflection of my own weaknesses—"most people aren't as narcissistic as you!" I thought. But I have slowly become convinced. Maybe not by the total number of people, but certainly by the total volume of tweets, far more than half of the people you see on Twitter are like the waiter in Hollywood who wants to tell you about his screenplay. I think we all know of this Rosetta stone at some level, but partly out of good manners (often an enemy of social criticism) we keep our awareness submerged.

To reiterate, Twitter is an audition held in a panopticon and there's probably no one on the other side of the glass

This has a couple of effects. Firstly, people write in a way which is implicitly dissatisfied with its real audience, and aimed at an audience cooler and more popular than their actual readers. There's something eerie about the fact that—at least to a degree- *they are not talking to you.*

Secondly, people have sensibly realized that because it's a very big audition they've got to stand out. Just being very good and incisive as a writer is *not* a great strategy for standing out. Too many people are playing that game—worse, too many people playing

that game who also have something else to offer—e.g. good looks or a compelling life story. So if you want to win you've got to find a niche and that niche probably can't just be "is a good writer" unless perhaps you are very good indeed, but a lot more people think they are exceptional than are.

As a further consequence of point two, people are more hostile because it's an easy way to stand out.

Thirdly, interactions have a subtly strategic relationship quality. This is most obvious when people tweet things like "favourite this for a compliment" or "I reciprocate follows", but those seem to me to be just the most obvious manifestations. This air of "you help me stand out and I'll help you stand out" is pervasive.

But fourthly, and above all, the joy of activity that is *within itself* -that does not point to some greater ambition—is tapered and in some cases eliminated.

You've probably already clocked that each of these four facets makes human interaction less authentic.

And all this for the sake of an internal mental observer that, as a general rule, doesn't reflect anything real.

---

(1) We are also fearing that someone might be watching and swoop down to cancel us, but that's a discussion for another time. To briefly elaborate—not only are we seeking good publicity, we're also afraid of the bad. The end result is that we're doing all the work of celebrities with few of the rewards. Though not quite "none of the rewards" because there is a certain pleasure to the celebrity LARPing thatTwitter and related offer.

The paradox of high expectations: The more you demand, the less you get

In various fairly common situations, demanding more can result in receiving less.

I.

The kinds of situation I am talking about are ubiquitous, but we'll start with employee hiring. Suppose you are running a job search, and are primarily interested in some desirable talent T. Perhaps T is years of experience using some application or programming language.

Now suppose that people come in two varieties, liars and truth tellers. Truth tellers will truthfully (or mostly truthfully) state their level of T. Liars will claim to have whatever level of T they think will make you most likely to hire them. Suppose further that every truth-teller who meets your expectations will apply, and every single liar, regardless of whether they meet your expectations, will apply. Assume the average T value is the same for liars and truth-tellers.

Let's say the population=1100 and 100 of them are liars. Now suppose you set the required level of T at 2 standard deviations above the mean. Assuming a normal distribution, 25 truth tellers will apply, and 100 liars. If you have no way of telling liars apart from truth-tellers, you have an 80% chance of hiring a liar, if you pick from among the applicants who meet the threshold at random.

Meanwhile if you'd set the level of T you were demanding at 1 standard deviation above the mean, you'd have a less than 40% chance of hiring a liar assuming you pick from among the applicants who meet the threshold at random.

If, for example, the value of a candidate to you is equal to their T-value expressed in standard deviations from the mean (positive or negative) you will actually get a lower T score on average by setting the cutoff at T=2, than if you'd just set it at 1.

A lot of readers are probably thinking that our simplifying assumption that employers are no better than chance at spotting liars is too harsh. I have a few things to say to that, viz:

A) The processes most employers use to determine quality past the resume stage—referee checks and interviews—are in many ways easier to navigate if you're a little bit loose with the truth. It would actually not surprise me if the real process filters out more truth-tellers than liars.

B) Available empirical evidence suggests that people are generally totally delusional about their ability to spot liars, and spotting liars is actually incredibly difficult.

C) Even if employers have some skill in filtering liars, if that skill is less than complete, it remains true that, past a certain point, increasing your expectations simply makes it more likely you will get a liar.

One interesting sub-case here is where you are interested in multiple traits, some of which you can check in applicants more easily than others. In these circumstances it may pay to set relatively modest minimum thresholds for the traits you cannot easily check, but adopt a policy of "more=better" for traits that you can accurately check.

II.

Perhaps the most extreme, and comical, variant of this phenomena companies that demand candidates who have more years of experience in programming languages than those languages existed. I don't know if this has ever actually happened, or if it is just a persistent urban legend, but any company that tries this is guaranteed to receive liars.

Although job searches are among the most obvious areas in which this paradox arises, it comes up in other areas. For example, seeking quotes, choosing between products and any other domain where the phrase "too good to be true" comes to mind.

When we consider that a pattern of high demands might turn once truth-tellers into liars, something like this phenomena could explain the flourishing of hypocrisy in some morally rigorous communities. Past a certain point, lying becomes a more workable strategy than actually trying to better yourself.

Demand the very best and you might get much, much less.

Postscript:

One reader suggests that reasoning similar to this is why you should find scientific papers just short of significance considerably more trustworthy.

Another commenter on Reddit suggests that we leave out one very important specification—a third type of person he calls an exaggerator. An exaggerator does not lie to an indefinite degree, but instead exaggerates their qualifications by a fixed amount. If such people are common, it may be prudent to demand a T value of 2, if your real minimum is 1.

This is a really great point, and I guess the overall picture here is that it's difficult to know in the abstract what kinds of demands and criteria are the most likely to work. We've certainly demonstrated that under some conditions, demanding more will lead to less in expectation. In other conditions though this doubtless isn't true. If there is an overall takeaway here, it is that the matter is very complex and unintended consequences abound. I don't really have any algorithmic advice, but it is probably worth thinking through a variety of possibilities and balancing different concerns against each other.

Movements are always a distorted lens on the ideas they embody

Please don't hate me for this one. I don't think I'm better than other people. I'm just laying out in blunt terms what others have tiptoed around.

1. **The problem**

I want to spell something out in this post that I think many people know, but which goes under articulated because it's kind of mean. I apologise for this. I really do feel poorly about it. I don't like measuring people up in this way, but it's one of those things you've got to be sharp about to be honest about.

I remember reading an internal socialist party document once. It was very careful in its words, but essentially the document argued that during periods of "downturn" in class struggle socialist organisations which were "keeping the flame going" would inevitably attract dead-enders, no-hopers, people who were mostly looking for a social outlet etc.

In the bible (1 Corinthians), St Paul remarks of the Christians of his time: "Brothers, consider the time of your calling: Not many of you were wise by human standards; not many were powerful; not many were of noble birth. But God chose the foolish things of the world to shame the wise; God chose the weak things of the world to shame the strong" i.e., the people joining up to Christianity at the time *even by their own*

admission were not people conventionally seen as intelligent, financially successful or strong.

This isn't exclusive to tiny religious and political sects. It applies to pretty much all social movements. I don't want to be mean, so I won't name the movement but there is a certain social movement which I have in mind that regularly makes people remark "Urrrgh, X sounds cool in theory, but have you met the Xers? No thank you."

Bluntly, movements always have and will attract a lot of dead-enders. This is because normal people don't need movements as an outlet—they already have friends and are capable of curating their own social lives. Dead enders are more likely to have trouble in this department, thus are more likely to gain something from the constructed social life offered by movements.

2. **N.B.**

Most of the examples I use here are from political movements because that's where I've spent a lot of my life. However I believe the dynamics I describe apply to a lot of other movements, from music (Punk) to even some philosophical movements (e.g. New Atheism)

3. **The taxonomy**

There are essentially four types of people who join relatively marginal social, religious and political movements, as well as certain types of intellectual and artistic movements:

1. People who aren't doing so well in the game of life.
2. Thoughtful, whip-smart, highly committed and charismatic people.
3. People intermediate between 1 & 2 with some features of both.
4. Genuinely normal people.

Category 4 is the rarest—so rare that we won't talk about it much further. Category 2 is the second rarest, more or less invariably. Whether 1 or 3 is the most common probably depends on what the movement is about and the context. I myself probably started life as a 1 and clawed my way, through the infinite patience of mentors, to a 3.

Why are these types particularly attracted to social movements? Well, it's no great mystery.

Category 1—The losers join social movements because:

A) They're lonely, and a social movement is a group of people who sort of have to be friends with you.

B) When you're behind anyway, why not gamble on something that might make it big? Become an early adopter of a plan to change the world?

Or the more sincere reason: **C)** Because their beliefs and values have been formed through a lifetime of exclusion from the dominant power structures. Thus their values and the values of critical social movements often have much in common.

To be clear, the reason the "losers" can't curate their own social lives is not always a lack of social skills. Sometimes, for example, they are perfectly charismatic but have burnt all their bridges through intermittent erratic behaviour. Sometimes they've just been unlucky in some way. A loser in this sense is someone who has turned to a social movement *because they have to in order to cope*. There's more than one path to that outcome.

Category 2—The stars are attracted to social movements for one of two reasons depending on the individual and how cynical they are

A) because it allows them to be a big fish in a small pond

B) because they're deeply committed to their beliefs, and will pursue them even if means hanging out with uncool people.

Category 3—The inbetweeners are attracted to these social movements for both sets of reasons in varying degrees.

4. **The resultant neuroses**

People in these groups, especially those who aren't dead-enders, are acutely aware of the dead-ender problem. The number of people who have started or led movements for a time, only to lament that they can't stand their own followers, is huge. Often these laments are some variation on "they only understood the *form* of what I was teaching, not the *essence*." Rightly or wrongly, the dead enders are accused of not getting the ideas, or at least not the deeper underlying truths of those ideas. I don't know if it is actually true that the dead-enders don't really get the ideas, but it's a common perception. Certainly, some forms of dead-enderism seem to correlate with a lack of insight.

Sometimes people even go so far as to claim that the majority of people following the movement are actually a block on its success—that the movement would be better off leaner, but higher quality.

My favourite contemporary example of sniping at one's own followers is the grandees of dirtbag left Twitter who are constantly complaining that their "reply guys" just ape them by repeating phrases like "normal country" and "Hellworld" without understanding the spirit of critical irony and convention busting that was meant to power it. Of course, this being Twitter & the Dirtbag Left, it's hard to be sure how much of the scorn is performative irony and how much is real exasperation—but my guess is "all of it is both ironic and heartfelt at the same time". There's something very funny about seeing would-be underground rebels reduced to complaining about slavish followers.

As has been remarked by 50 million other authors, the tension between the leaders and followers is most especially a problem for movements based around "breaking the rules". The leaders watch in despair as their modes of rule-breaking become the new rule by people *who just don't get what it's really about*.

I suspect these neuroses have been around for a lot longer than we realise. Histories of movements are mostly written about those we would classify as the "stars" of movements. Stars are also more likely to write the histories as well. This tends to submerge underneath the waters of Lethe lot of the angst about dead-enders. Still, nothing is entirely lost. It's been a while since I've read any of it, but I remember getting the impression reading between the lines that some of these issues plagued 19th and early 20th-century socialism.

5. **The contradictory role of size**

Now you might think that the dynamics that I've described here only apply to radical ideas. They don't. I've met the youth wings and gone to branch meetings of all 2 1/2 major parties in my country *if anything they were even a little bit sadder than the socialist groups, the libertarian groups etc.*

Despite the fact that seemingly more mainstream movements are no less filled with losers, it does seem that as movements expand the people in them get more normal, and as they shrink they get less so. I remember Occupy in my home town, which continued for an unusually long period of time. It very clearly moved back and forth between these extremes(1).

David Graeber makes a similar point about anarchist and direct action groups in *Direct Action, an Ethnography*. He also points out that the proportion of women in these groups rises as the group grows larger, often closer in time to an important "action" or

"event". I to have observed this, and I believe this gender dynamic to be linked, although I'm not going to feign a hypothesis about why it is so.

In *Ruling the Void* Peter Mair talks about the hollowing out of political civil society. Parties all over the world have fewer and fewer active participants. Based on accounts I have read of social movements throughout history, I do not think that this is isolated to formal political parties. Especially from 1990 to the GFC, activity was minimal. After the GFC there has been something of a slow recovery of activism, but it remains to be seen how permanent and significant it is. Activity in various movements is still modest compared with the 70's. This is not only true of political movements—even cultural movements seem pallid now. Sometimes it feels like only strange weirdos like myself remain.

6. **The implications: evaluation**

So why am I laying this out? Well, over the last five years or so, I've noticed a rise in arguments of the form "X is a bad idea because the people who follow X aren't living normal fulfilled lives". This takes a lot of forms, e.g. duelling Chad vs Virgin memes etc.

Now I don't think we can dismiss this as an argument by simply labelling it ad hominem. If an idea is causing problems in living, and/or turning out and attracting unbalanced people, we should at least understand why before moving on. It is prima facie evidence that an idea doesn't work in practice if everyone who likes it isn't doing so well.

But I would caution against taking this argument too seriously, precisely because movements, especially their zealots, *are always like this*, and now maybe more so than ever.

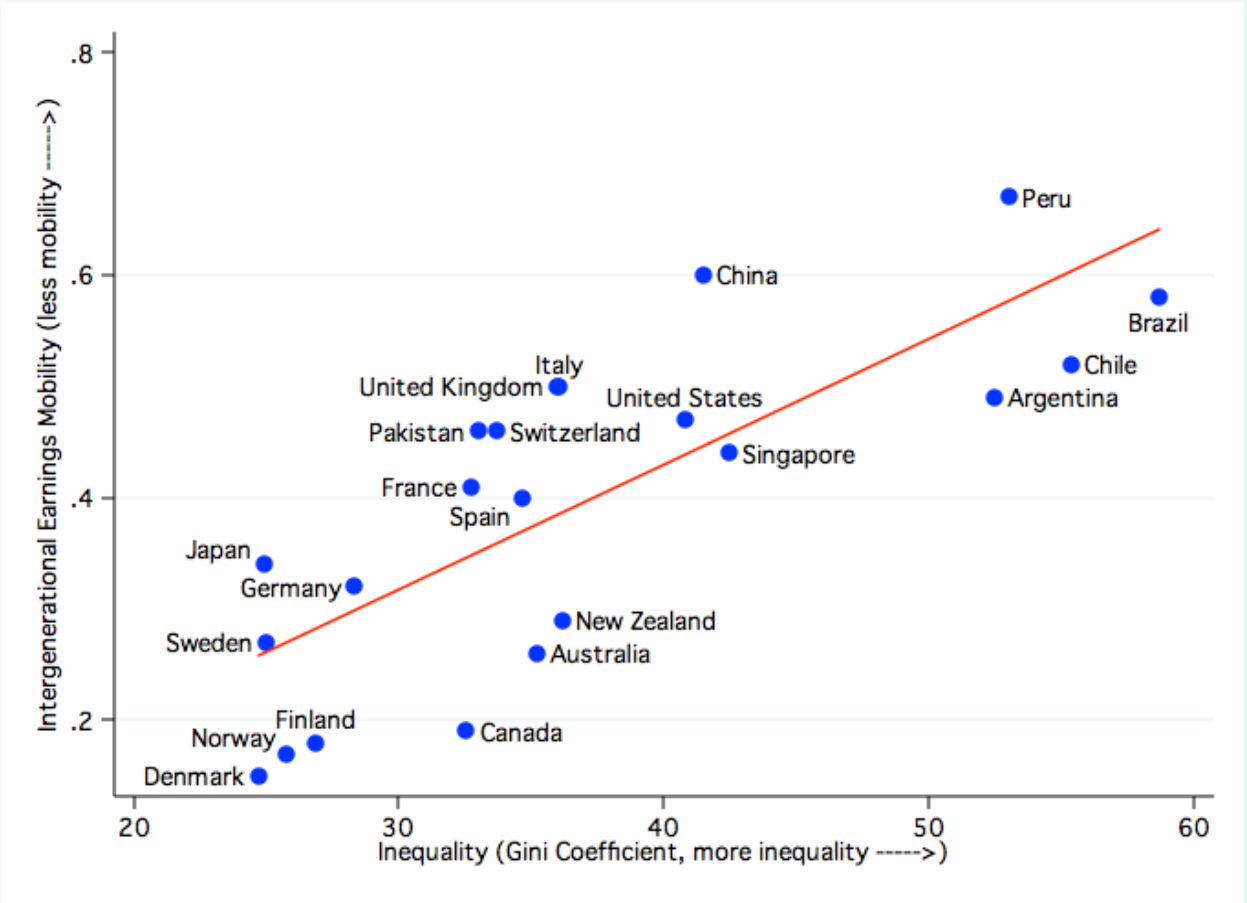
To finish on an upbeat note, maybe, in the end, the 'losers' have the last laugh, because in getting involved, in making ideas and arguments, even if they are never the leaders, they are more likely than the conventionally successful people who stay away from movements to affect history.

Notes: on Michael Sandel's "The Tyranny of Meritocracy"

Section I

- **It's obvious once it's pointed out, but it really is amazing how few people realise "anyone can succeed in America" is a big fuck you to everyone who hasn't succeeded in America. Some of those people vote!**
- **Sandel has gone through Trump's speeches and noticed that there are few, if any, sentiments along the line of "anyone can succeed in America " or** "America is a land of opportunity" whereas there are hundreds of such statements in the words of his predecessors. He thinks this is not a coincidence, and I am inclined to agree.
- **Sandel is right that a very interesting feature of political rhetoric is that it makes statements about the essence of how things are as a call for change– paradoxically**. He gives the example of Obama saying words to the effect of "America is a land of equal opportunity, therefore we need to give every child a fair child a fair chance at an education". When you step back for a moment the two halves of this statement are completely contradictory! Yet in political rhetoric nothing could be more natural than to merge them. This trope of presenting a *change* as really a reflection of an underlying pre-existing *essence* is very old. C.f. Spartan political debates where people would interminably argue that their reforms were really restoring things to their constitutional essence.
- **There are three political strategies for selling ideas and ideologies to people unhappy with their place in an unequal system**, viz: **A) The treasure your wins strategy**– tell people that they should think instead about being happy to have beaten out those they lapped in the race of life, and get them on the side of maintaining inequality that way. **B) The reroll the dice strategy**. Say "Yes, you're right, there are some injustices in the system, so let's rework the rules to be a little fairer and then you, or at least your children, can reroll the dice—you might win this time! **C) The gap reduction strategy** reduces the size of the gap between winners and losers.

- **At first glimpse it might look like A=Conservatism, B=Liberalism, C=Leftism. But it's not quite that simple—at least not all the time.** Conservatives often offer a chance at rerolling the dice—"you'll be able to compete as a small business owner once big government corporatism is gone". Liberals sometimes go for a `treasure your wins strategy—"those rubes want to devalue your hardwon education—don't let them, experts like you deserve to run the country". Leftists mostly focus on C—gap reduction, but do talk about B—rerolling the dice—sometimes.
- **Although many people have heard of the Great Gatsby curve by now, it's always worth a reminder that the dilemma between "equality of outcome" and "equality of opportunity" is fake because the two are deeply correlated.** :



- **I've been thinking a bit about decadence lately—especially thoughts occasioned by Ibn Khaldun, Peter Turchin and my good friend Kieran Latty.** I think one window into understanding what most people get wrong about decadence is the concept of "luxury". People rightly

associate luxury with decadence *but for the wrong reasons*. The reason luxury is associated with decadence is **not** because it's decadent to have nice things—at least in any meaningful conception of decadence. No, the defining feature of luxury is not having pretty or scrumptious or fragrant things—it's conspicuous consumption—trying to outshine your neighbours through your purchases. Decadence is a state of affairs wherein people -elites to be specific—view their primary goal as competing within society, rather than trying to advance society. I think this is Ibn Khaldun's sense of decadence—and he is right that it does destroy kingdoms and empires. This is also Peter Turchin's understanding of decline, and he is likely right that it is caused, in part, by an overproduction of elites.

- **History isn't going to rap you on the knuckles because people are having a bit too much gay sex or men have long hair now, history is going to rap you on the knuckles if people aren't committed to larger projects than themselves.** One of the main symptoms of that is luxury—people start buying nice clothes not because they are nice, but because they are nicer than yours. One of the great disservices the right has done us is tainting the concept of decadence—a very important concept—and making it merely a vehicle of bigotry and trad aesthetics.
- **One way to understand this is in terms of a Marxist theory of the state re: America right now** yes I promise this will sweep back round to Sandel. So the US is doing very poorly—low growth rates, much unrest etc and a lot of it is clearly a result of bad governance. The Marxist theory of the state suggests that the state is the steering committee of the capitalists as a whole, and this is to my mind, largely true. Only that steering committee is meant to provide a *synthesis* of those interests. Obviously there will be conflicts between industries, but the state is meant to rise above that, at least "in the main". However, the US state is failing to do that. It's not an articulated whole of corporate interests—it's a sack full of them – many contradictory—stuffed in greedily. That's political decadence, and it is linked in manifold ways to individual decadence.
- **Sandel's point can be understood as—meritocracy is the ideology of decadent elites, so obsessed with their internecine competitions that they have forgotten their obligations to the common good and to the weak**, so puffed up on their little games that they think fairness within their little club—and fairness in the rules of admission to their club—is the big game of governance.
- **I want to end this section by reiterating that none of this means that either Sandel or I don't think every child should have a chance to succeed**. On the contrary, if you care about that, paradoxically, stepping back and focusing instead on making sure everyone has decent living

conditions might be the best way to achieve your end, because the empirical evidence shows there is a strong correlation between low economic inequality and equality of opportunity. Meritocracy abstracted from this broader social project is—at best—about putting guardrails down to ensure that the children of quasi-elites have a fair shot at becoming full elites.

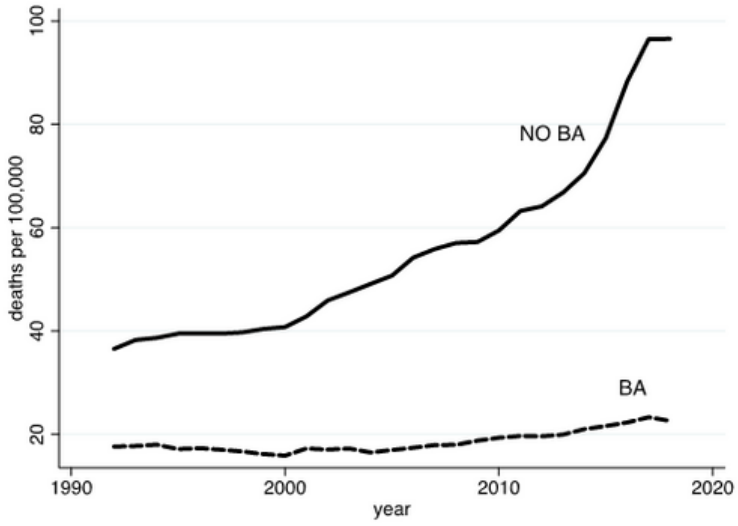
Section II

- A reader asked of the last post "what exactly do you and Sandel mean by this "meritocracy" you are critiquing? You've said you're not opposed to the idea that the best person for the job should generally get that job, so in what sense are you not meritocrats?" That's a great question. Sandel doesn't quite spell out what he means by meritocracy, but I think there are a few separable but related theses, viz: **A)** Meritocracy as a theory of desert—people deserve remuneration corresponding to the degree of contribution their talent allows **B)** A theory of political priorities—the most important thing is to ensure "equality of opportunity". Making sure that people at the bottom have a decent life is of secondary importance (if even that). **C)** A mode of rhetoric, focusing on equality of opportunity, the virtues of rule by the smart etc. **D)** A willingness to concentrate honor and dignity in hands of the "meritorious" e.g. talented. **E)** A theory of how political problems arise, viz not having the "best people" with the "smartest ideas" take care of them—instead of conflicts of value, practical interests and moral beliefs. Hence meritocrats find themselves committed to what I and Scott Alexander have previously called **"Mistake theory"** and contrasted with "conflict theory".
- Sandel makes a big deal out of what he calls "smart language", especially the language of "smart ideas" and "smart policies." Such language is attractive during a period of political polarization where debates over what is *right* can seem so much more intractable than debates over what is *clever*, but the solution is a false one. Calling your policies the smart ones insults the intelligence of your interlocutor. What started as an attempt to avoid rancour feeds into it. This is a great point about mistake theories of political conflict generally—in trying to avoid conflict they risk inflaming it by implying that one's opponents just aren't as clever!
- Through analysis of rhetoric Sandel argues that the Obama era was a great time for mistake theory (although he doesn't use the term of course). However, Sandel also makes the point that if mistake theory was dominant during the Obama era as a mode of rhetoric and form of ideology that doesn't mean it was any truer as

an analysis of the political conditions. Rather, constantly talking about "solving problems" "commonsense solutions" "smart solutions" etc. may have predominated *precisely because politics was a morass of endless bickering at the time*. The mistake theory rhetoric was wishful thinking and/or a futile effort at peace making.

- One of Sandel's more insightful points is: the obvious unfairness of unmeritocratic societies at least gave a "handle" that social critics and popular movements could grasp onto in an attempt to fight for a better world and better conditions for the lower classes. Critics could say "you have more than me and that's arbitary, give me more" and there could be a conflict over that demand. Meritocratic societies on the other hand are "frictionless" in a way which doesn't dissipate popular anger, but instead leaves it inchoate—and potentially more destructive. Just because the apparent "fairness" of the system compared with overt aristocracy *(not that the system really is fair, even on its own terms)* makes articulating anger harder, doesn't make the anger go away.
- One of the most important points of the book is that people are, now more than ever before, concerned as much with the distribution of honor as with resources. This is something that, sad to say, I think the left has often gotten wrong historically. People are as worried about getting their lives to fit a meaningful narrative in which they matter as they are about making sure they'll always have food to put on the table. This may seem like a very obvious point—and surely on some level we all know it—yet I must admit that I've often failed to fully get it. Articulating a form of historical materialism which is fully alive to this need is important.
- Sandel discusses the history of meritocracy at length. Two things that stood out especially to me in his discussion. 1. Meritocracy may have been a cold war innovation—a desperate society turning away from entrenched privilege to ensure the best and the brightest would be in place to fight communism. 2. A president of Harvard who was one of the original champions of meritocracy, and insisted it was distinct from equality of outcome, nonetheless couldn't prevent massive favoritism towards legacy admissions precisely because the alumni were rich and powerful enough to get their own way. This is a superb example of how massive inequality of outcome will tend to eat away at equality of opportunity—whatever noble intentions to keep them separate.
- Sandel reminds us of Erica Scharrer's fascinating studies of inept men in sticoms. Over time there has been a tendency to portray men as more foolish—bumbling etc. in sitcoms, and women as competent and holding the family together. However the impact has fallen unevenly, with working class fathers much more likely to be portrayed as incompetent and useless and the tendency has been increasing overtime. This, suggests Sandel in conjunction with a variety of other evidence, is part of a pervasive cultural denigration of all those without a college

degree, but perhaps especially men without a college degree. Sandel even suggests that the massive spike in deaths of despair among people—especially men—without a college degree may be linked to this general cultural denigration.



Deaths of despair (suicides, drug overdoses, and alcoholic liver disease), ages 25-64, 1992 to 2018, for those with and without a college degree (age-adjusted)

- This ties into something I've been thinking about for a long time. One argument we often see on the left is that it's okay to casually tease men, engage in joking (and even not so joking) misandry etc. because men aren't an oppressed group. Now there's a lot of truth to this, but it neglects another truth—when you attack a group, the brunt of that attack will fall on the weakest and most marginalized members of that group. Rich and powerful men will laugh off criticism of men in general. The people who get hit will be poorer, lower status men. The same is true of attempts to "discipline" the bad behavior of men or other dominant groups. The more powerful members of the group will often evade discipline, and it will instead fall upon the less powerful members—poorer, more likely to be disabled etc. A more sophisticated approach to social structure is required!
- An extreme example of how "castigation of the privileged" can harm the vulnerable—those people who thoughtlessly say things like "I can't understand white men who still end up homeless, you had everything going for you and you still failed". A disgraceful sentiment.

- Sandel argues there has been too much emphasis placed on *distributive* justice, when really we should be equally interested in *contributive justice*. Everyone wants to feel like they are making a contribution to society. At least at the margin, people's most pressing unfulfilled desires are often not about consumption, but about feeling like they are making a valuable contribution to society. A lot of why inequality stings is not because it means we can consume less, it's because society is quantifiably scorning our contribution. An ethic of competitive meritocracy doesn't make the losers feel like their contribution is very significant. Politicians, economists and political philosophers alike have been guilty of making people's identities as consumers primary over their identity as producers.
- When I think about my own greatest fantasies—to be an acclaimed writer, singer or philosopher, to be a hero, it's notable that they are all fantasies not of *taking* from society but of *giving* to society and of being recognised for that contribution. I don't think I am unusual in this regard.
- It's worth noting that there are resources within my broad intellectual tradition—Marxism—for recognising and addressing exactly this point. The idea of the producer alienated from his product in any number of ways is just as fundamental to Marxism as the idea of material scarcity.
- Contributive justice might be all very good and well as a goal now, but let's say that AI gets better and better and consequently the value of many people's labour falls. How can we aspire to give everyone contributive justice under those conditions? Sandel doesn't grapple with this problem, but I think it's an interesting one. Let's say that the transhuman solution of "upgrading" everybody so that they can make a material contribution isn't viable—at least for a time.
- I think under these conditions the best we could probably do would be to encourage people to see themselves as contributing *through actions and ways of being that are inherently meaningful*. Joy, friendship, self-discovery, making art. To shift from contribution through the production of extrinsically valuable goods to contribution through the """"production""""" of intrinsically valuable goods. Another good that people can provide that doesn't necessarily depend on skilfulness is giving their own preferences about what is ultimately, non-instrumentally good in democratic deliberation. I'm not saying it will be easy, but I think there could be a path to give people a sense of making a meaningful contribution even in a post-scarcity society.
- One thought of Sandel's that will stick with me—the writing is in the wall for neoliberalism as currently understood. Even its most ardent supporters should be able to recognise this by now. The question then is not *will the present "mode" of capitalism fall apart, but what will replace it?* Authoritarian centralism? Quasi-feudalism? A replay of the post-war years

with renewed unions? War and barbarism? Literal fascism? Social democracy? Socialism? I don't know if the future is open, but it is certainly unknown. All we know is that the present won't last. Understanding that the tower is going to fall, we just don't know which way yet, is an important shift in perspective.

PART 3: OBSESSIONS AND COMPULSIONS

Harm OCD, a brief introduction

This article contains discussions of very disturbing topics including mental illness, sexual assault, racism, pedophilia, sexism, extreme violence, etc. I speak about this stuff in brutal detail, because I think it's important we be clear on the nature of this illness and the kinds of distress it can create.

A reader asked me to explain more about my experience with OCD. I declined because I've already written a fair bit about it and feared I didn't have anything else interesting to say. However, some events on Twitter—the horrific spectacle of some faux-woke people saying to the mentally ill that they are just bad people, convinced me that an explainer about a form of mental illness called Harm OCD is necessary.

At several points, I help myself to immoderate language. I want to be clear that this isn't because I'm being melodramatic, it's because I feel extremely immoderately about this topic.

1.

OCD is a mental illness affecting 1-2% of the population. While it's not a competition, and I would never downgrade anyone else's suffering, there are signs that OCD can be among the most serious of anxiety disorders. For example, it typically must be treated with much higher doses of SSRIs than other forms of anxiety and depression. OCD is not a fun or cute disorder.

Wikipedia defines it quite adequately: *"Obsessive-compulsive disorder (OCD) is a mental disorder in which a person has certain thoughts repeatedly (called "obsessions") or feels the need to perform certain routines repeatedly (called "compulsions") to an extent that generates distress or impairs general functioning. The person is unable to control either the thoughts or activities for more than a short period of time."*

Harm OCD is a form of OCD. In harm OCD the patient falsely fears that they will, or even that they have already in the past, do something destructive and unethical, opposed to their own values. Psychiatrists and psychologists have documented numerous instances of what we are calling harm OCD. Examples of harm OCD form an important part of diagnostic implements like the Y-BOCS. However, the impetus for thinking of it as its own unique category of OCD, with special complications and features, has largely been driven by the community of its sufferers.

To be clear, harm OCD is not a separate illness from other forms of OCD, it's only a different kind of manifestation. It's not uncommon for people suffering harm OCD to have experienced other types of obsessions and compulsions in the past, or even to experience them at the same time.

Some examples of harm OCD include:

- Susie is concerned that when she is cooking for her children she will include bleach in the dish and kill them.
- William is worried that he will sleepwalk in the night and stab his partner to death.
- Jiang is worried that she will grope and sexually assault a child while walking down the street. She fears she may have already done this and not remembered.
- Morgan fears that he will, absentmindedly "on autopilot", write out a confession to a horrific crime he hasn't committed. He never throws out any paper in case he has accidentally written confessions on them.
- David is worried that he will just "flip out", "lose control" and start beating up people as he walks down the street.
- Kiara is worried that she will mutter the N-word when she is on the phone with a black client under her breath.

- Lisa is concerned she will accidentally perform a hit and run. Every time she hits a bump in the road she reverses, to make sure it wasn't her running over someone and not realizing it.
- Trevon repeatedly has images pop into his head of him raping and murdering people. He worries this means he is a psychopath who wants to rape and murder people.
- Ayaan had sex with a woman several years ago. When he looks back on it he worries *what if the woman wasn't really consenting*? What if he pressured her into it, not meaning to? He replays tiny details of the event in his mind.

An attentive reader might be thinking "I see the obsessions, but where are the compulsions' '. Compulsions in this form of OCD can be subtle and less obvious than, say, cleaning compulsions. A sufferer might keep their hands in their pocket so that they don't accidentally grope someone, or mentally replay the events of the day, looking for evidence that they hurt someone inattentively. They could try to think good thoughts, to cancel out the bad thoughts. They may mentally rehearse arguments about whether some occasion on which they had sex was coercive or not. Forms of OCD where the compulsions are very well hidden are sometimes called Pure O OCD. Some people even believe it is possible to have OCD with no compulsions at all, even hidden ones—just obsessions—however this is controversial.

In addition to harm OCD there are two forms of OCD that are closely related and worth mentioning because they also integrally involve the values and identity of the subject.

One is sexuality OCD, where the patient is afraid that they have a different sexuality than the one they usually regard themselves as having. Some examples, A) a straight man worries that he is secretly a homosexual because he keeps having unwanted thoughts about having sex with men pop into his head. B) A man or woman who keeps having unwanted images of having sex with children come into their head and worries this means they are a pedophile.

The other is scrupulosity OCD, where the sufferer worries about blaspheming against religion, or not being sincere in their faith, etc., etc. They may have, say, involuntary sexual thoughts about the virgin Mary and be deeply distressed by this. Martin Luther may have triggered the reformation, partly because he was driven to distress by scrupulosity.

People with any form of harm OCD are not dangerous. Professor Ross Menzies once told me words to the following effect:

If hypothetically he wanted to be absolutely sure that he was not associating with someone who would perpetuate a hit and run, he would be well advised to spend time with an OCD sufferer afraid of doing just that. If he wanted to be absolutely sure his children would not be abused, he would leave them in the care of a person with pedophile OCD. If he wanted to associate with someone who was controlled and unlikely to lash out, he would associate with someone whose OCD made them terrified of exploding in anger, and so on.

He even suggested that harm OCD could be thought of as the opposite end of the spectrum to sociopathy, and disorders where there is an absence or mutedness of conscience.

The tragedy of harm OCD is that, through a process akin to natural selection, the mind picks the worst possible fear—the one most opposed to the sufferer's deepest values—to inflict on them. This is what sticks the best—what is hardest to shake and what will keep popping into the head. The sufferer is effectively being punished for the depth of their opposition to running people over, or raping children, or being racist, or stabbing their partner.

2.

Twice now I have seen rubbish on Twitter of the following form:

1. Someone says they have intrusive thoughts about performing a deviant act [the two examples I've seen: sexually assaulting a child, saying the N-word].
2. People pile on and say that it's disgusting that a person would admit to being tempted to such acts.

This is a horrific thing to do, and contemptible among those of them that should know better. It is one of the most hurtful things you could possibly do to a person with harm OCD, and their special vulnerabilities and fears mean that it really is no better than physical violence against them. If you only take one thing from this article, let it be this **having an intrusive thought about doing x does not mean you are tempted to do x.** Quite the opposite, it means that you find x especially abhorrent or terrifying. Many people with harm OCD may *believe* that they are tempted to do horrific things but they are not.

How can a person falsely believe that they are tempted to do something horrific? Well, there are many ways, but here is one example. Try not to think of a pink elephant. You can't do it, right? Well, imagine if instead the thought you were trying to block out was serially raping and murdering women. You can't force yourself not to think about something. In the end, the thought pops into your head so many times that you begin to fear that it must be because you desire it. In truth, it's the exact opposite. Because you want nothing more in the world than not to imagine these horrific thoughts you can't stop thinking about them.

It should be obvious why telling a person with harm OCD that their condition proves they are dangerous or evil is one of the cruelest and awful things you can do to a person with words.

3.

Harm OCD is a self-concealing illness.

Because it specifically involves *things you are horrified of* it makes you not want to talk about it. You fear that if others find out about these thoughts, you may be ostracised. Tragically, as the events on Twitter, I mentioned show, this can sometimes be true.

I believe that the real incidence of people with harm OCD is probably much higher than we know, for this reason.

Finally, people who talk about their harmful OCD in detail, despite it representing everything they abhor, are heroes. They are risking their reputations and their mental stability to tell you something very painful about their experience. May God have mercy on your soul if you use that as an opportunity to turn on them.

Everywhere you go, you always take the weather with you

Everywhere you go you always take the weather with you

-Crowded House "Weather With You".

In this essay I want to talk about philosophy and OCD. Not for the purpose of creating a "philosophy of OCD", but rather to find ways in which the vantage point of OCD can aid philosophical inquiry.

0. An obsessive compulsive life

Obsessive compulsive disorder is a horrible thing and I would not wish it on anyone. That said, living with OCD has given me a certain way of seeing the world. The insights and viewpoints of that way of seeing are not unique to OCD, but OCD is certainly one path to them. I decided to write this to share what the refining fire of a lifetime with OCD has given me.

1. **Generalities**

[I have removed this section as a basic definition of OCD is covered elsewhere in this anthology- in the previous essay on Harm OCD]

2. My specifics

Anyone who doesn't have a guilty conscience needs one, and anyone who does have one, doesn't need one.

–Something I once read on a Magic the Gathering Card

One more bit of general background—the kind of OCD I suffer from.

I suffer from a form of OCD that makes me afraid of enacting physical, sexual or psychological violence on others, sometimes called "Harm OCD". I fear both the harm I might do to others, and the consequences I might face as a result of doing such harm. Although this form of OCD is quite widely understood among experts, not many people in the general public are aware of it. Partly I think this is because of the nature of the content of this form of OCD—it's self-censoring. Although it's embarrassing, you might be willing to admit that you're afraid of catching aids in a public bathroom for example. However, admitting that you're afraid of sexually assaulting someone, or kicking a child in the face as you walk down the street, is scarier. Thus knowledge of this form of OCD hasn't seeped into the public consciousness.

A lot, but not all of my fears, centre on the following nexus. Within this nexus there are numerous variations on a theme and elaborations, and I am perhaps misrepresenting what is going on in some respects, but this is the basic story. Suppose a random thought of groping someone on the subway pops into your head. At first, so long as you recognise that they are not really *your thoughts or desires* this thought may not be very frightening. Suppose though that you began to worry that because such thoughts had popped into your head so many times, that you might perform these actions on autopilot—mechanical reflex—without even being aware of it, let alone intending it. You now have a plausible story about why you should be afraid of these thoughts, and also a story about why being afraid of these thoughts is dangerous in itself. Trying to be rid of these thoughts becomes like trying not to think of a pink elephant. You will think these thoughts all the more because you perceive them as dangerous, and in your mind that makes them even more dangerous.

Here is a simplified schemata of the process:

Thought of harming someone>>>Thought that you've had that thought so many times you might do it on autopilot>>>The thought is now perceived as

dangerous>>>This makes you want to stop thinking about it>>>This makes you think about doing it more>>>Loop continues

Because you're worried about doing these things 'automatically' you don't even have the comfort of reasoning 'well nothing has happened yet and it's been a long time, so it should be okay'. Part of what you're afraid of is that you might be doing such things all the time, and just not noticing. Perhaps the victims were perhaps too shocked or scared to say anything.

The irony of harm OCD is that all available evidence suggests that people with it are unusually appalled by violence, and thus much less likely to commit it than the general population. There are no recorded cases of someone with OCD committing violence unintentionally, or in a way in which relates to their fears.

3. The Red Queen Hypothesis of OCD: Fear as genetic algorithm

"Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!"

-The Red Queen, *Through the Looking Glass*

One of the things I find interesting about OCD is the process by which it generates fearful obsessions perfectly targeted against its subjects- expertly finding their weaknesses. In coming to understand this, I realised that OCD is a self-generated mental parasite, operating like an evolutionary algorithm.

Fearful thoughts are generated randomly, those fears that are successful (in the sense of capturing attention and thus cognitive resources to elaborate on them) are selected—just as with a genetic algorithm. These fears give rise to lineages of similar fears and variations on a theme which are in turn elaborated on. Over time, a thoughtful sufferer will come to see holes in their fears, and this leads to arms race where the fears evolve in response to keep up with the growing dialogue between fears and defences, like the red queen theory of evolution in biology, or, like the runaway evolutionary processes of cancer, constantly subverting and being subverted by our natural defences.

Coming to understand that a part of me—my fearful thoughts—was evolving and acting with logic that was blind to, and actually opposed to my own purposes was terrifying, like finding an eldritch horror in your own breast. However, it was also a breakthrough not only in my struggle with OCD, but in my struggle to know myself.

4. You don't really know scepticism until you've fought delusions

You usually learn facts through indirect sources, and you know that these sources can be faulty, even if you consider that unlikely. On the other hand the basis of your reasoning about, and inferences from, those facts is immediately present to you. If you don't find an inference convincing, you will most likely drop it. Thus we have a cognitive blindspot when it comes to being sceptical about our own reasoning—doubting the premises on which we build our inferences is easy, but doubting your own reasoning—the inferences from those premises—is hard, because *it wouldn't be your reasoning if you weren't already directly convinced it was persuasive*. Thus there is an asymmetry in our capacity to doubt our premises and our inferential moves- the former is easier than the latter.

Your accepting a sequence of deductive inferences as valid is partly based on your inability to imagine ways it might be wrong. Even in the case of inductive reasoning, your accepting a sequence of inferences as valid is at least based on it being hard to imagine a *plausible* scenario in which the premises hold but the conclusions don't. On the other hand, imagining a factual premise is wrong is easy—"I misremembered" or "Maybe someone gave me wrong information". This asymmetry is a real shame, because as many errors are made through faulty inferences as through faulty premises, so we should be just as sceptical of our reasoning as our premises.

A few concrete examples. Witness the incredulity, confusion and even anger which can arise when someone is corrected about the Monty Hall problem. Another example is the way in which many scientists worry about whether their methodology is generating reliable data, but then often scarcely worry at all about whether there are any gaps in their method of statistical analysis.

I believe OCD can help with this—and maybe even if you don't have OCD reflecting on the experience of an otherwise reasonable person with OCD can shake false confidence in your inferential capacities. Having been in a state where you feel like everything is about to come crashing down on you again and again, only to find each time that your carefully manufactured chains of inferences about what would happen next meant nothing *even though you could see nothing wrong with them* is very educational. In my experience OCD typically takes a series of commonplace facts and real observations and weaves a story which seems perfectly plausible to the sufferer according to which the only reasonable inference from these facts is upcoming disaster. The inferences involved *seem* very persuasive, but turn out to be meaningless—so you have to become good at recognising that there is no link between how persuasive it seems to infer D from A, B, & C, and the real validity of that inference.

Recognising that an individual chain of reasoning can always be invalid—no matter how good it looks—is the first step to becoming a hedgehog. More on that later.

5. Epistemic judo—turning doubt in upon itself

"Doubt your doubts…"

-Switchfoot

I began to articulate the idea of epistemic judo. In real Judo, you leverage your opponent's weight against them. In epistemic judo, the person afflicted with doubt and inconfidence leverages those feelings against their own anxieties. Empirical evidence suggests that people with OCD are unusually sceptical and doubtful, even about things not related to their fears. Doubt and the trait of being doubtful seems to be intimately related to the aetiology and processes of OCD, leading some to call OCD the disease of doubt. How could the OCD sufferer leverage this capacity to doubt everything against their own fears?

The first step is to think of your mind as a mechanism for generating understanding of the world. Return to our earlier model of OCD as an evolutionary algorithm for finding blindspots in your epistemic processes, and using them to trigger fear. You can try to beat it by seeing why every single lie it tells you isn't plausible after all, and this will help to some extent, but to make progress past a certain point *you have to stop believing that just because a fear seems plausible, it is plausible—even if you haven't got a great counter argument for it yet*. If you can't take this step, your OCD will keep throwing up new fears and variations on the old fears faster than you can spot what is wrong with them. In other words, you have to stop believing that your own OCD driven cognitions are reliable, even if you can't see the specifics of why they are wrong. You may not have found the logical gap in the argument that you should be afraid of X yet, but eventually you will, and even if you don't, chances are that it exists.

The strategy to achieve this is to turn the sceptical/doubting mindset OCD generates against itself, until you reflexively doubt your fear arguments. To a certain extent, you have to stop believing in your own ability to reason about topics that overlap with your OCD. Ordinary people have trouble engaging in blithe and automatic scepticism of plausible sounding ideas they themselves have generated. However, people with OCD have so much practice questioning and doubting everything, what is it to doubt one more thing? Eventually I turned "I am not qualified to think about these topics well or impartially" into a mantra.

6. Arguing with your feelings

"Thoughts aren't facts"

-Common Parable

"Feelings aren't facts"

-Another common parable

Most narratives and most bits of folk wisdom, tell us to trust our feelings—our "intuition". Perhaps the hardest part of grappling with OCD, and one of the most educational, is learning **not** to trust your intuition. Sometimes feelings are right and sometimes they are wrong. Just because you feel like **SOMETHING IS VERY WRONG** doesn't necessarily mean shit. If you want to overcome OCD, you have to learn to be skeptical of- and sometimes even outright ignore- your own intuition.

If you think of feelings as unlike thoughts—as beyond debate, let alone volition—you will find this difficult. Hence you must come to understand the similarities between thoughts and feelings. Both are rooted in ideas about how the world is, and neither are immune from criticism. For example, being "afraid" of a spider is almost always accompanied by the thought that it can hurt you, or crawl all over you in an unpleasant way. You feel the way you do because you have a certain model of the world. There is no sharp distinction between feelings and thoughts.

The wrong way to think about this is to think of it in terms of feelings and emotions being always inferior to formal reasoning—of feelings as inferior or flawed thoughts. Actually, some feelings are very accurate. Sometimes the explicit chains of reasoning we call 'thinking' are much worse than the pattern matching we think of as feelings or intuitions, and sometimes the opposite is true. This is unsurprising because, again, emotion was never discontinuous with reason, and both feeling and reasoning are subject to critique, revision and scrutiny.

People have accused me of valorising the Cartesian thinking subject at the expense of the embodied subject of feminist epistemologies yada yada with this view, but I think the opposite is true. Only by challenging the false separation of thought and feeling can we see what is wrong with both the mystical valorisation of intuition and the obnoxious assertion that one has transcended reason for pure rational assessment. "Trust your

feelings" and "Suppress your feelings so you can be truly rational" are two sides of the coin, for they both posit a separation, and disagree only on which has priority.

7. Becoming a hedgehog: OCD and the struggle against single model thinking

"Homo unius libri timeo"

-St Thomas Aquinas

Almost all OCD type fears have the following structure—a prediction about how the future will be (or in the case of guilt, how the past was) based on a specific sequence of events in a causal pattern. At least one of A-Z must happen and then this will follow this and then either this or that will happen but then in either case necessarily this must precede… and so on. Even though every individual step might sound plausible, something almost always breaks down, and so the fears of people with OCD are almost never true. This form of unitary model thinking is an extreme type of the "Hedgehog" cognitive style identified by Isiah Berlin, and contrasted with the "Fox" cognitive style, based on a more flexible appraisal of different factors and plausibilities. Another name for these modes of thought is Euclidean versus Babylonian methodology.

There's a great deal of evidence that the hedgehog type approach breaks down not just in OCD, but in things like expert attempts to predict future events. Overall being a hedgehog isn't a great strategy (although beware anyone who tells you it is **always** wrong—we wouldn't want to be meta-hedgehogs).

As I saw how being a hedgehog was making me mentally ill, I made an effort to become more of a fox, not just with regards to OCD, but throughout all of my life and activities. For example I tried to view the challenge of predicting the future not as a matter of charting a sequence of events, like falling dominoes or a Rube-Goldberg machine, but instead tallying power factors, tendencies, resources, general drifts of various coalitions and past trends, and throwing in a generous dollop of pattern matching.

8. I am not special: Taking the outside view on yourself

There's a kind of soft narcissism that OCD forces on you. It doesn't help that OCD makes you live inside your own head and spend so much time talking to yourself . Your problems seem very distinct from the problems of other people who have your

conditions. Their fears sound so absurd, whereas your own sound so reasonable to you (due to having been specially crafted for *your* cognitive blindspots).

It can seem like the standard treatments could never help—e.g.: dammit I don't need to relax, I need to establish that I won't grope someone! Or: What's the point in going to a psychologist—I'm just as clever as they are, any point they can make about how my fears are illogical I will have certainly thought of myself! Yet all of this turns out to be false, chances are you're really not so different from other people, and will benefit from, and be harmed by, roughly the same things as them. Certainly that was my experience.

Recognising that you have all the fallibility—and strength—of other people is incredibly liberating. It's okay to reason as follows:

"Hey, this stuff seems plausible to me, but I am a delusional person, so rather than using what the base odds seem to me to be, I should think about it from an outside point of view. How many people with OCD feel their OCD fears are plausible (almost all), and how many actually turn out to be right (almost none, and **literally none** in cases of harm OCD like you have). Therefore it's massively unlikely that you will be the first person with harm OCD ever to have your fears realised. The fact that you think you are an exception to this rule, doesn't matter a damn, because there is a 100% chance you would think this, given your OCD."

In other words, it's okay to take the outside view on yourself. For example, assessing the plausibility of your own thoughts just like a clinician who was only vaguely aware of the content of your thoughts, but knew your diagnosis, would assess them.

This isn't just about the plausibility of your fear arguments. You have to take the outside view on a lot of lifestyle and treatment factors. For example, "While it seems that my problem is the dangerous situation I'm in, so sleep won't help, if I actually get some rest, the experience of countless humans over many millennia indicates the situation will probably seem very different". Or "While I don't feel consciously lonely, and socialising doesn't feel like a priority since I believe my life is about to collapse, the experience of the human race as a whole would suggest it's a bad thing that I haven't talked to anyone

except my parents in almost three days and that this may be the real reason for my terror. I need to go meet up with a friend".

I've taken this attitude and applied it to other areas of my life. I don't regard my studies and writings as attempts to find truth in any personal capacity because the odds that I will discover something important are much lower than the odds that my novel ideas are just crankery. Instead my hope is that I will contribute to the social process of truth seeking. I look upon my own work indulgently—from the outside perspective it is true that my work is likely air and puffery, but also from the outside perspective it's true that you could say that of just about anyone so it is important that we don't let the fact that our work will almost certainly be either wrong or unoriginal stop us.

9. OCD and non-self

"I may venture to affirm of the rest of mankind, that they are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement."

-David Hume

It's kind of pat, but it's got to be said. At a certain point, gazing at the changes within yourself, and understanding them as intertwining and sometimes contending forces you begin to see the inside of your own brain not so much as a monarchy with an ego-king, but as a republic, full of traitors, excitable mobs and hard working but beleaguered bureaucrats. Seeing yourself as a balance of forces can inspire vertigo, but it can also inspire hope. A unitary subject with attributes attached is hard to change, a contending swarm can shift.

Perhaps the best metaphor is the self as an ecology. Particular thoughts and feelings you have are like interacting species lineages, each fighting for survival. Drives are like the terrain. Perceptions are like the weather, carrying in nutrients for the organisms and taking them away. You are this jungle- this tangle of evolutionary game theory.

Less abstractly, I remember \sitting in a lecture when I was much younger, shaking with terror that I could scarcely conceal. I began to imagine myself as like a boat on a stormy sea, except because I was both the boat and the sea I could never escape. Wherever I

went I would always take the weather with me. As painful as this moment was, it stayed with me because it was the beginning of insight.

Later, I beheld the mirror and saw nothing reflected there, and so realised that “I” wasn’t beholding the mirror at all, I was the mirror, I was the room in front of it, and I was the wall it stood on. There is no person in that room, and there could be no person in it, because that room is a series of components which make a person.

Obsessive Compulsive Disorder and the origins of religion

Obsessive Compulsive Disorder (OCD) is well known to interact with pre-existing cultural and religious beliefs and practices. I was interested in coming at it from the other direction—is there evidence that culture, particularly religious culture, has been shaped by OCD?

Although I can’t prove it absolutely, I believe there is enough evidence to make the idea intriguing, and worthy of further study by someone better suited than me. The idea that mental illness might be linked to the origins of religion is not new, but I have rarely seen the idea in the context of OCD specifically. However, I suspect that there is a connection, and the effect has been profound.

It is worth stating categorically that a link between religion and OCD is not intended to denigrate religion, nor OCD sufferers. As a lifelong sufferer of OCD myself, such an intention is the furthest thing from my mind. OCD is associated with many positive traits and traits which can be positive in some situations—caution, concern for our impact on others, cleanliness, seeking certainty, aiming to systematise etc. Although these traits tend to get out of control in OCD, at other times and parts of life they can be profoundly useful. If there is a relationship between OCD and the development of religion, the contribution of OCD could well be positive.

Nor does saying that OCD played a role in the formation of religion imply that great religious figures all had OCD (though many may have). Rather, ambient forms of OCD might have influenced religious thinkers and their interpreters. One might also believe that even people without clinical OCD can tap into an OCD like mindset, especially when concerns as vast as salvation and divinity are on the line.

Personally I see the link between OCD and religion as one of OCD and certain aspects of religion utilising the same ‘circuitry’, although outright cases of OCD most likely play a role in the development of religion as well.

The idea of a link between religion and OCD is far older than my poor thoughts (notably, Sigmund Freud championed it, although there is surprisingly little development in later literature), but I wanted to explore it in a more contemporary way. I've gone looking for information on people who have tried to develop a connection, but as best I can determine, most of the literature tends to be practical and epidemiological, with little of it coming from a studies-of-religion perspective.

What follows is my attempt to trace what I see as shared themes between OCD and religion. The exact role, if any, that OCD, and OCD-like mental states, played in the formation of specific religious beliefs is probably in most cases lost to time now, yet as we will see there may be grounds to suspect it was a rather large one.

Themes that OCD might have contributed to the development of religion:

In what follows we examine five parallels between OCD and religious thinking, and speculate about the possibility that OCD-like modes of thinking may have contributed to their development in religion. These five: 1. Concerns about purity 2. Fear of offending the sacred 3. Rituals and ritualistic behaviour 4. complex, detailed and guilt driven forms moral reasoning, and 5. ideas of demonic or external possession. While many of the parallels we discuss are not entirely unique to obsessive compulsive disorder, the overall picture is of a striking resonance between the focused but fearful reasoning of OCD, and similar trends within religion.

1. **OCD and similar mental states as an explanation for the origin of religious fears around contamination and purity**

A common concept in OCD is the fear of contamination, although today this fear is often expressed in terms of the germ theory of disease, prior to the development of such it would likely have taken different forms.

Religious categories of ritual and dietary purity are often worked out with a logic that is at once rigorous and dreamlike, mimicking the 'style' of OCD thinking. This overlap in content (concern about purity) and style of thinking (internally coherent, deeply elaborate yet bizarre) is striking. There is a deep concern for how various unclean things might infect various clean things, of what is capable of 'carrying' or 'transmitting' uncleanness, about making sharp rulings in unusual and borderline cases and so on.

The book says do not mix a kid goat in its mother's milk. Easy enough. But the book says don't do it three times? There must be something extra here. Better not mix meat and milk, and keep a four hour gap between eating one and the other, just in case.

2. **Fear of offending the sacred, blasphemy, scrupulosity and the origins of sacredness**

A very common form of OCD is scrupulosity, a fear of offending the divine through uncontrollable sacrilegious thoughts, words, mental images etc, or through very small lapses in action or words that would seem unimportant to many. At first glance one would assume that the idea of scrupulosity presupposes the prior existence of a concept of the sacred, yet I can imagine ways in which symptoms of scrupulosity and a sense of the absolute sacred might co-develop. The belief that one must not even think—let alone say—certain things about a being, even in jest or illustration, elevates it to a higher level of sacredness. For sacredness to exist, the possibility of blasphemy must exist.

The constant guilty struggle with one's own thoughts and micro-behaviours about the sacred can further drive greater religious engagement, and the development of cycles of repentance, perhaps contributing to the often dramatic rituals of self-abasing penance and the purging of sin common to so many religions.

3. **OCD and the origins of religious ritual and ritual magic**

Numerous OCD behaviours are rituals, in the sense of stereotyped series of actions intended to influence seemingly unrelated actions. Tapping a spoon three times to make sure your husband does not die on a ship, for example. Or Saying the Lord's Prayer 27 times to make sure that your children do not die of a horrible disease.

It is easy to see how, in a society which did not necessarily dismiss such rituals like ours tends to, someone might develop a belief about a necessary series of actions and this might be adopted by the society at large, especially if the ritual 'worked' several times.

OCD even comes with a ready-made explanation of why rituals sometimes fail, since there is an overwhelming concern that one will slightly mess up one of the innumerable complex parts. Ritual failure is itself an important theme in many religions.

4. **OCD, guilt and the development of categories of moral reasoning**

A common form of OCD is a supreme concern that one will, or that one already has, acted in a way wholly dissonant with one's moral values. One form in particular is the fear of negligently or actively causing others harm, or violating the moral order of the universe, with actions that would, to most outside observers, look harmless.

Replaying past actions for evidence that one did the wrong thing in one's mind and endlessly vetting possible actions, could drive the creation of new categories used in moral and religious thinking.

To choose a few examples, categories of Islamic jurisprudence, Catholic moral philosophy and—perhaps above all—the Jewish Halakha—display an exemplary combination of deep caution, deep concern, deep subtlety and exceptional systematisation. Relatively harmless or even completely harmless behaviours are condemned in the strictest language. St Thomas Aquinas claimed that masturbation was worse than murder, Jesus said that thinking about another woman sexually was a form of adultery and that calling one's brother a fool was worthy of the fires of hell, the Talmud says that it is better to be burned alive than to embarrass someone in public. The fusion of intense anxieties with careful legalistic thinking bears a weird resemblance to OCD, and the longer one reads the texts, the stronger the parallels seem.

Neurological evidence suggests that OCD may be a disease of hypermoralism, and is deeply implicated in the over activity of neural pathways associated with moral reasoning and thinking (forming a possible inverse to antisocial personality disorder). It seems to me very plausible that inclinations towards hypermoralism may well be one of the motors driving the development of moral thought in a religious context.

But unreasonable rules give rise to resistance, and religion is no exception. Resistance to hypermoralism, by OCD sufferers who had become sick of it might form another motor driving moral development within religious traditions. An ongoing dialectic, sometimes creative and sometimes destructive, between the spirit and the letter of the law, is a theme in every religious tradition I have ever studied. I'm not just trying to have it both ways here—there's a least one very good case study of this kind of 'bending in the opposite way' reaction to hypermoralism by an OCD sufferer with a profound effect on the history of religion: Martin Luther, leader of the protestant reformation. Martin Luther is fairly well established to have suffered from scrupulosity, a religious form of OCD. Luther's continual sufferings, and deep fear that his behaviour was inadequate almost certainly contributed to the theological views which caused him to propose that belief in god alone was enough for salvation, thus attempting to cut the cycle of fearful hypermoralism.

5. OCD and demonic possession

While many mental illnesses can be seen as 'demonic possession' by those without a psychological background, the potential for OCD to be perceived in this way is often forgotten. While I don't think the majority of cases of 'demonic possession' are really OCD in disguise, I suspect a substantial minority may be.

I would argue that many forms of OCD—such as uncontrollable thoughts of the violent, sexual or sacrilegious which go against the values of the sufferer could be seen in some societies as a sign of mental demonic harassment. Many sufferers of OCD (falsely) believe that they are experiencing barely controllable urges to do violence. Such feelings could very easily be mistaken for a sign of demonic activity, either by the sufferer, or by society at large.

6. Religious (Self)-Consciousness and internal struggle

But more importantly than extreme cases like demon possession there is the form of self-consciousness associated with many (though not all) forms of religious life, the sense of oneself as a tempted being—a being who is constantly in an unwilling dialogue with dark forces that wish one to sin. Whether these be conceived of as as internal, external (like a demonic tempter) or something ambiguous and intermediate (like the Yetzer Hara) the effect is the same, the self is seen as divided, and torn between obedience and sin.

The parallel in OCD is found mostly in cases of morally charged OCD—scrupulosity, harm obsessions and sexuality obsessions. The sufferer can be confronted with a sense of division about themselves, about who one is, what one has done and what one might do—a kind of fragmentary consciousness in which parts of one's consciousness sometimes seem work with a vicious semblance of autonomy in order to perversely thwart the whole is fundamental. Many sufferers of OCD begin to conceptualise their OCD thoughts and impulses as a cruel and bizarre stranger living in their mind.

We might also point to the internal struggle between doubt and belief that plays such a role in the life of the believer, and such a role in OCD. Certainly there's a parallel in concepts here, although whether there's a deeper or historically important parallel is harder to say.

Summing up the historical case

You're a psychologist, someone comes to you and pitches a hypothetical. There'a a patient who keeps strict rules regarding bathing, refuses to mix certain foods, repeats certain words at certain precise intervals throughout the day, regards certain behaviours as absolutely taboo for reasons which they either cannot articulate, or which seem bizarre, will only dress in certain specific ways, is afraid that they are a deeply wicked person and struggles with questions about their moral identity.

If you had to take a guess at the diagnosis, what would it be?

Evidence of an association today

To bolster our historical associations, let us take a brief look at the very large literature suggesting that religion and OCD remain linked today. Degree of religiosity is linked to likelihood of developing OCD, this has been validated for a wide variety of monotheistic religions including:

Judaism,

Christianity

and Islam.

While formal research has focused on these religions (largely because of their popularity in the regions where research has been conducted) the internet is full of heartrending stories of people struggling with these problems in all kinds of religious contexts including Buddhism, Hinduism and Jainism.

There is little doubt that membership in at least some religious communities is associated with OCD. The direction of causation remains unclear.

Conclusions

What we've presented here is very partial and tentative. The idea of a strong connection is far from certain, but many individually modest pieces of evidence can form a greater image.

I want to reiterate that it is not my intention to contribute to a crass reductionism here. We are not saying that major religious figures were all OCD sufferers. We're not making an argument that religion is a disease. Reasoning of the form "OCD might have contributed to religion, OCD is a pathology, therefore religion is pathological" is lazy and dangerous.

A more fruitful approach, and what I'd like to see, is greater investigation of the link between OCD, mental states which parallel OCD, and the development and survival of religions over time.

Rather than denigrate religion, I would hope such a study might deepen our appreciation of the social role of both religion and obsessive compulsive disorder.

A final quotation:

"Luther's first years in the priory were thus a time of interior tension, spiritual struggle and suffering. The hopeless feeling that he was not numbered among the elect but among the reprobate overcame him and grew stronger as he grew more and more conscious that he did not fulfil God's commandments in all things. Since he began early to condemn as sin every movement of natural appetite, even though unwilling, and since, with his exuberant vitality, such movements kept recurring, he supposed himself to be full of sin, and no prayer, fasting or confession could free him of this terror."

Karl Adam, *Roots of the Reformation*, as quoted here, itself an invaluable resource.

Lessons I squeezed from a lifelong severe mental illness

In the past I've written both an essay on the philosophical aspects of what I learnt from OCD and another essay on the practical measures that helped me keep it under control. Now I wanted to convey the little life lessons I learnt from decades of chronic mental pain. Stuff that isn't directly about the disease, but might help anyone.

Plot friendship

Friends have kept me alive. I have rather a lot of them, despite not being that charismatic (or perhaps that's false modesty and I am charismatic, but if so, it was very hard won). You need friends. Chances are you will not get friends organically unless you are in school or *maybe* in college. So you must make it happen. That means taking deliberate and systematic action.

Do these things 1. Give yourself chances to meet people. 2. Keep a list of the people you want to know better. 3. Make a list of the friends you do have and keep an eye out that you're not neglecting any of them.

I know this sounds weird, maybe even creepy. However, we live in a newly atomised world—it was much less atomised even half a lifetime ago. The conventional wisdom hasn't kept up with the atomisation. A much more active approach to your social life is needed than TV and books would indicate. If that makes you feel a little bit like the talented Mr Ripley so be it. Plot friendship.

You can't enjoy what you grasp

You can't really live if you hold your life too dear. So many things I worried about losing through my fears—freedom, dignity, life, reputation—I could not enjoy because I spent so long fearing for them. If you hold onto things too tightly, you lose time with them just as surely as if you had thrown them away.

Music helps

One morning I felt so bad—trapped and cold. Then I turned on the Mystery of Love by Sufjan Stevens and I felt completely fine.

We all know this right? "Music can help with depression!" but it's more than that. Music, poetry, visual art—we tend to treat these things as "nice". In reality, these are load bearing pillars of our soul. You are cross-stitched with little bits of art and science and philosophy. That song lyric you can't stop thinking about? It's integrated into you now. Engineers and artists are the two closest things to sorcerers that exist. We don't take them anywhere seriously enough. Cherish those who give you the materials to build yourself.

I regret every time I'm scoffed at someone singing along to music. I recently took up singing and can't believe I didn't do so years ago. Yousician costs 40 dollars a month and takes 10 minutes a day. Maybe my efforts are laughable, but it connects me with strips and veins of myself that would otherwise be inaccessible.

Life is better with a crazy ambition

I have a crazy dream. I won't say "working towards that dream keeps me sane", but it's one of the little things that helped a bit. If you asked me what the odds of me pulling it off were I'd say about 1 in 30,000. That it is technically possible and that I want it to happen are about the only two things in its favour. I'm not going to reveal what it is—it's too embarrassing—just think of some juvenile quest and you'll be in the right vicinity. I selected it on the basis that even if it fails it will help me achieve a bunch of valuable intermediate goals along the way. Consider picking one for yourself. Make sure it's not dangerous to you or anyone else. If you're not embarrassed by the thought of sharing it publicly it's probably not crazy enough. Go wild, I give you permission to be a megalomaniac about this.

Whatever you do to the least sets the standards for everyone else.

My anxieties revolve around the fear that I might do something terrible. As such, I have spent far more time than most people contemplating what it would be like to be one of the most reviled people in society. The conclusion I keep coming back to is that the way we treat the most reviled people—think convicted serial rapists or killers for example—sets the minimum standard for everyone, and that has a trickle up effect on those nearish the bottom. It is in everyone's interest to set that standard higher, because none of us know what the future holds. It may be true that you will never be in quite such a reviled position, and it may be true that right now you are far from the bottom. However, anyone can fall low enough that the minimum standard will be one of your few safety nets. You could end up like one of the people in Ronson's book, for example.

Prayer helps even if you don't believe

I don't believe in God. I wish I did, but that's not something I can control. Nonetheless when I become depressed or anxious I often pray. Sometimes I pray for myself, but mostly I pray for some kind of universal redemption or outpouring of mercy. Sometimes I weep while I pray. Sometimes I feel angry, and ask him for answers like Job shouting into the whirlwind. Sometimes I protest like a Karen asking to speak with the manager of the universe, appalled at what the customer service has done to 100 billion humans. Then I reflect that if God- omnipotent, omnibenevolent- is real, then by definition he has his reasons. I have no idea why, but this process helps me.

I dreamt I saw

A portion of God's face

Only by his grace

Did I not see more

Lest I end like the sea

Ends a tributary creek

Rich in the mystery of darkness

Abundant in the revelation of light

The fire of his eyes is justice

And from that flame rises

The incense of mercy

And all the heat is love

And the cold is love too,

Lest we burn away

All there is, is love

Please.

My method for dealing with anxiety

I wanted to outline the methods, strategies and approaches that have helped me with anxiety (moderate to severe OCD). Although I hope you might find something useful here, you may not. Seek qualified and competent psychological advice. To make it very clear that this is only my experience, and not general advice, I've written a lot of the post in the first person. I apologise if that gets grating.

The critical concepts are:

A). *Insight* . You have insight when you regard your problem as anxiety and you recognise, on an intellectual if not an emotional level, that your anxieties are either very likely unreal, or at least greatly overplayed. When you lack insight you regard your problem as the content of your anxieties. Insight exists on a spectrum.

B). *Rumination*. To ruminate is to spend excessive amounts of time thinking through your anxieties.

C). *Attention shifting*. The antidote to rumination is attention shifting. As a strategy, attention shifting means a decisive, insistent and active refusal to engage with or entertain anxiety thoughts.

The aim is to raise the degree of insight while avoiding rumination. You might be wondering how this is possible—how can both happen at once? after all, to challenge a thought, doesn't one have to think about it? The strategy that has worked for me is multi-headed. When I face a spike in anxiety, I employ the following steps:

1. Get my general factors right: All the usual advice, plus a few me-specific things. If I am not taking a med, start one. If I am taking a med have a conversation about whether I should raise the dose with my doctor. Consider supplements with some scientific basis, such as ashwagandha. Exercise! An anxiety beating combo for me is walking while listening to podcasts. Cut stressors and utilise supports. Keep in mind there's a delicate balance here—I don't want to cut stressful activities in an avoidant way, or give into anxiety and become a hermit. I also need to remind myself to use the supports I have available.

2. Conceptualise myself as ill and recognise the limitations of my own reasoning: This is going to sound paradoxical or self-limiting to a lot of people, but I've found it useful. I get myself into a headspace where I am very aware that I have an anxiety disorder, and that having an anxiety disorder means that some of my thinking will be irrational and distorted even when it "feels" right. I develop and use a mantras that will help me remember the limitations of my own thinking. For example: "Anxiety is distorting my thinking so I can't trust my own judgements about my fears".

3. Read some articles by experts about my specific form of anxiety: It is sometimes also wise to read the stories of people who have similar anxieties to myself (there are plenty of forums where people post descriptions of what they are worried about). Reading experts and fellow sufferers writing about fears like mine is an

extension of point 2. The aim is to begin to see my thoughts as the manifestation of a common illness by reading about similar manifestations in other people—similar content to their worries. The aim is to create a perspective shift in which I see myself as a sick person, rather than as a person endangered by the content of my worries.

N.B. This could very easily become a weird kind of compulsion—a sort of reassurance seeking—although that hasn't been my experience thus far. This is important to monitor.

4. Meditate on the optimistic meta-induction (article here): Really burn it into my mind. This step also builds on point 2.

All these steps help reduce degree of belief in the content of fears and worries, without going down the rabbit hole of rumination.

5. Go see a psychologist ASAP: Even in a session where I don't make much progress (and that is rare), seeing a psychologist is at least a tangible reminder that I am a person with an anxiety problem, and not a person with a "whatever I am worrying about" problem—so once again it comes back to insight.

6. Create a defined, delimited space in which I am permitted to ponder my anxieties: About 15 minutes a day is appropriate. I sometimes even want to limit the time to "When I see my psychologist". In that space, called "worry time", it is appropriate to ponder my anxieties. When thoughts occur outside that space, I push them aside, and mentally acknowledge that the right time to think about those worries is in worry time. It's easier to dismiss things if there is a space where those things are permitted, so I am not simply saying "No, bad thought!" I am instead saying "In its proper time".

7. In worry time, bookend every chain of reasoning with the thought that, whatever my assesment of the situation good or bad, the truth is probably better because anxiety is systematically biasing my thinking towards bad: Weave this liberally through my thoughts.

8. If I find myself worrying or ruminating outside worry time, squash those thoughts with prejudice by shifting my attention: Think about something else. This is easiest when I act quickly. Often there is a moment of vague anxiety just before I think my first anxious thought in a train of worries—I try to use that moment to squash my anxieties and force myself to think about something different. In our culture there is a bias against the idea of squashing thoughts. It's seen as a form of cowardice. In truth, going over the same thoughts over and over again is a futile search for comfort and certainty—abandoning that takes courage. Novels and pop psychology teach us that if we can only find some perfect key, we might find some great revelation or answer to our worry. In truth, when we leave thoughts that are no longer constructive behind, we create the space for personal growth.

9. Aim to think everything through once and once only: In practice the way I do this is oriented around sessions with my psychologist. I prepare before the session (making notes to bring in to the session) and then really go through every argument and fact I can think of relevant to an issue with my psych in detail before to a conclusion. If I am later plagued by some aspect of that same topic again again, I consider whether it's really novel, or just a rehash. If it's really a novel perspective or fact, I resolve to mention it in my next session with my psychologist, or at least not think about it till worry time. If it's just a rehash, I remind myself that I've already made a determination, under the closest to ideal and rational conditions I can, and that whatever my feelings, re-litigating it can only lead me further away from the truth (in expectation).

OCD, mental illness and "cancel culture"

I want to use OCD as a lens to understand and critique what is sometimes called cancel culture.

Fear of cancel culture is moving like an infection through the OCD sufferers. Why, and what does this tell us? Someone posted an interesting thread on the OCD subreddit recently:

"Does anyone else have an OCD-fuelled fear of being "cancelled"?

Throwaway because this is a subject I'm mega anxious about!!!

Basically, the title: Does anyone else have really distressing obsessive thoughts about being "cancelled" online/losing their livelihood etc. because of stupid things they did as a teen? I'm just petrified that people will find out about bad stuff from my past, they'll tell my employer, I'll lose my job and never find one again, my life will be ruined, that kind of thing. When the thoughts hit they just cycle and cycle around my brain and I find it so hard to function because what's the point in doing anything if my life will eventually be ruined for things I regret doing? I have to go through all of these compulsive behaviors like Googling my name to see if anything bad comes up and seeking reassurance from loved ones just to feel slightly better, and even then the fear still remains a bit.

Am I totally alone in feeling this way, or have other people experienced it to? If so, how did you manage to overcome it?"

Some interesting comments on the thread:

"I've even stopped pursuing my dream of being a published author because I'm scared that it'll only lead to being "cancelled"..."

"i'm still pursuing that same dream but i'm gonna use a pen name."

“Cancel Cultures themselves have probably made the same or similar mistakes as the person that they are canceling. Either they have forgotten about it or kept it to themselves.”

“I’m so happy I found somebody who understands this I’m crying”

“ALL OF THE TIME!!! Seriously! I always ask my friends if anything I did back then and even now is awful or worthy of being cancelled, etc, and they look at me like I’m crazy.”

“One thing that helped me was realizing that people have had these kinds of concerns across cultures and eras and found them very disturbing. I don't know why, but seeing 'reputation' on this list of the eight worldly concerns really helped contextualize my own fear and put it into perspective (I am just one person feeling afraid about their reputation).”

“Very much. I’m waiting for someone to unveil the person I truly am underneath and for me to lose everything. Like sometimes it’s from things I know I’ve done and other times it’s just a vibe that the end is coming for me.”

There were the dozens of comments along the lines of “wow, were you in my brain”, “yes, 100% this” etc.

Finally, I saw this comment, which was so good I’m going to draw special attention to it:

I honestly believe that cancelling predates cancel culture... it's a part of a general belief system that people are disposable cogs in a machine that pervades our culture as a whole. Like in the past you could get 'cancelled' by major film studios for being gay, for example. It connects to the idea that if people are 'a problem' you can 'get rid of them' by firing them, locking them away, etc. to keep the core 'pure.' As OCD sufferers we're

disproportionately impacted by all forms of purity culture, which has deep roots in our society connected to racism, homophobia, and so on.

Then this caught my attention:

"Im a psychologist and this is becoming a super common theme for my clients"

So I did a bit of searching to see if it really does go beyond this. I found dozens of other threads about the fear of being cancelled on the OCD subreddit, many with dozens of commentators, picking one at random:

https://www.reddit.com/r/OCD/comments/hdup6b/cancelcallout_culture/

Then there were threads on many other forums

There were podcasts:

https://www.fearcastpodcast.com/2020/07/14/real-event-ocd/

And articles:

https://carleton.ca/determinants/2019/cancelled-overcoming-the-fear-of-a-social-media-presence-in-a-growing-call-out-culture/

https://adaa.org/learn-from-us/from-the-experts/blog-posts/consumer/metoo-latest-ocd-trigger

https://cognitivebehavioralstrategies.com/ocd-in-the-age-of-metoo-revisited/#.YHdvSegzZPY

This blog post by a sufferer is particularly good

https://notmakinglemonade.com/myblog/2020/2/5/im-so-ocd-about-scrupulosity

Etc.

So from research, It looks like a number of psychologists and psychiatrists are reporting a lot of OCD sufferers coming to their practices scared of being canceled, Metoo'd, or similar. Patients are also talking about it on the internet. I find this interesting and want to understand it, as a way of critiquing—and understanding—"cancel culture".

OCD is an opportunistic pathogen as one blogger noted. There are a lot of OCD sufferers paranoid rightnow about the coronavirus. This phenomenon of people being scared of cancellation because it's in the news isn't surprising in some ways, but I want to dig down into what it can further tell us.

I find the category of "cancel-culture" a little frustrating because it focuses attention on celebrities. The etymology here is telling "cancellation", coming from the idea of canceling a show, movie, book, or whatever due to a controversy. The old label of "callout culture" seems more adequate in that regard—less focused on big names and celebrities who can be, metaphorically and literally, pulled off the air. I care a lot less about some celebrity being canceled (although I do care) than I care about some poor nobody getting called out, screamed at, told to kill themselves, etc. on the internet.

But the move from "callout" to "cancel" is, in some ways, more accurate. "Cancellation" captures the killer instinct inherent in the phenomena. All pretense to a moral corrective in the term "callout" is now gone.

The other interesting thing about the lingo of "cancellation" is that an ordinary person who fears being canceled is implicitly comparing themselves to a celebrity. I think they are both very right, and very wrong to do so.

They are very right to compare themselves to a celebrity because the internet, particularly Twitter, Facebook and Instagram have turned fame into a matter of degree, rather than a qualitative distinction. I'm not famous, but if I desperately want 10,000 people to read something I write, I can make it happen. The line between "a big account" and ``internet famous'' is extremely blurry, and so is the line between internet famous

and real famous. This is not to say that there are important distinctions between famous and not here, but more than ever, they are quantitative rather than qualitative.

We are in the future that Andy Warhol imagined. We have our 15 minutes of fame, or at least the chance to strive for it. However, all that fame consists in is the bad bits of being a celebrity—the whiplashes in public opinion—with none of the cash or groupies, and with no agent to manage our PR.

Yet it is also very illogical to consider yourself like a celebrity in this sense. If you're an ordinary person, unless you get very unlucky (I'm not denying it's possible), no one cares if you said the N-word on a forum once when you were 15.

I'm not saying ordinary people are immune to cancellation. I once sat down to list everyone I personally knew personally IRL who had suffered a major reputational blow of some kind—from being accused of sexual assault to being exposed for some past act of racism. There were well over 30 people on that list. Some were pretty much as deserved as these things can ever be, others were tragic (an abuse victim accused of abuse by her abuser).

The truth, then, is complex. OCD, as it often does, has a cruel grain of rationality at its core. People are being cancelled. Somewhat randomly, sometimes with weak cause. The odds of it happening to you though, especially in a way that matters, are pretty small.

Who is vulnerable to shaming?

There's a quote I once saw on, of all places, a Magic: the Gathering card that has influenced my life ever since. "Those without a guilty conscience need one. Those with a guilty conscience don't". How does that apply here?

And just who is most affected by shaming online? I don't mean here just people who get explicitly "canceled"—the celebrities etc. I mean people who are, in one way or another,

humiliated, shamed, defamed, or otherwise caught up in the rough and tumble of online argument in which reputations are cheap.

Tautologically, the most harmed people by anything are the most vulnerable. The people most vulnerable to shaming come in a lot of varieties

Materially that is:

1. those who are easily fired or deprived of their livelihood.

2. those who without the money for a legal or PR team.

But also, emotionally and relevantly to OCD:

3.Those with pre-existing mental health problems or an anxious/depressive disposition and

4.Those with strong consciences.

An inevitable effect of shaming online in which, as one of the commentators above noted, people are treated as disposable cogs, is that it is not necessarily elites who are going to be driven out for their deviancy, but the weak, vulnerable and caring. All in the name of protecting the weak, vulnerable and caring.

One user's comment about how this stuff had gotten under their spiritual armor stood out to me:

Yep. I struggle with this. For me, it's more internal, though —I'm less worried about the potential effects of being "cancelled" (I do still worry about those as a creative —and also about the actual social interactions that would come with such a thing since I'm autisitic and have social anxiety, but those worries are less obsessive somehow) and more about "what if I really am a bad person?" It's sort of a metaphysical

contamination theme at it's core for me. I sometimes explain it as the bad parts of Twitter "cancel culture" living in my own head.

Now if my argument were that we should challenge cancel culture because it can be upsetting to people with OCD. That would be stupid. All sorts of things can be upsetting to people with OCD. My argument is a bit subtler than that. My argument is that the boiling-intense moral economies associated with these online spaces are more dangerous to psychologically and materially vulnerable people than to anyone else. OCD is just one example of this. Other examples include not just mental illnesses, but situations and conditions of precarity in all kinds and varieties.

Tu quoque

One of the reoccurring themes in that thread and others that I found interesting was "*why are these people doing this? Don't they know that they've surely done things that they could be canceled for as well?"* A while ago I did a survey on justice and mercy. One of the main hypotheses that I wanted to test was that people who had a negative appraisal of their own past would be less likely to engage in judgmental behavior online—or at least express the view that they should be more reluctant to engage in such behavior. Surprisingly I did not find such a link. Yet here was a thread full of people remarking that they think it's foolish to judge when we all have skeletons in our closets. Perhaps there is a link between aversion to judgment and a sense of guilt about one's own past, but only in certain kinds of mental illness? Much to ponder.

Regardless, I think we OCD sufferers have it right here. Judge not, lest when you are inevitably judged, your judges add insult to injury by also calling you a hypocrite.

En passant: "Cancel culture" and religion

At the moment there's a debate going on about whether that strange iteration of left-liberalism calling itself "social justice" can be considered a form of religion. I think I might write more about this later, but I wanted to make a few remarks.

As an atheist, it's hard for me to see it from the inside, but religions are often dialectics between *mercy* and *justice*. Such tug-of-wars have occurred many times in history, but intriguingly 2000 years ago a radical preacher from Nazareth came into conflict with the religious authorities of his day, in perhaps the most direct iteration of the clash between mercy and justice.

I'm not religious, and this isn't a religious, blog, but the words of Jesus here have interesting parallels with the present. He argued that in their tendency to emphasize strict adherence to the law over mercy, the existing religious authorities inevitably made hypocrites of themselves, because they preached rules that were too strict and comprehensive for themselves to keep, and preached them without exception. Thus they did in secret what they said not to do in public, corrupting their own virtue, even while they were concerned to demonstrate virtue. He compared them to tombs, bleached white and shining on the outside, but inside filled with rot and corruption.

Genuine concern for other people, he argued, was much, much harder to uphold than a list of laws. It demanded infinite concern and activity, rather than a box checking approach to goodness. It would require being honest about one's failings over respectability, and seeking forgiveness for those failings, even as we granted forgiveness to others. Genuine concern though, was more authentic than legalism. This is to say, in the dialectic of justice and mercy he championed mercy. In championing authenticity and direct concern over rules and habits, he was perhaps also the first existentialist. 2000 years later as we talk with each other about justice while trying to throttle each other on the internet I find that interesting. What would a Social Mercy Warrior look like?

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

PART 4: HOW DID THE LOVE OF WISDOM COME TO THIS?

Meeting Nietzsche at the limits of rationality and the limits of Analytic Philosophy

Note: I am not a Nietzsche Scholar. These are my musings and attempts to form an internal model of something. It is my hope that, even if they are wrong, they are wrong in an illustrative way. In this regard, let me appeal to Nietzsche himself: "It is certainly not the least charm of a theory that it is refutable; it is precisely thereby that it attracts the more subtle minds."

I've been reading Nietzsche lately. His genius is pretty clear to see, though on most things we disagree. My reading style is a little naive. I've picked up copies of his books, and I'm going through them by themselves, rather than reading them with secondary sources. So far I've finished *Beyond Good & Evil, The Birth of Tragedy, Thus Spake Zarathustra* and hope to go through more.

It's got me thinking about the limits of Analytic Philosophy. Nietzsche's own contempt of threads that would later weave together into Analytic Philosophy—British Empiricism, British Utilitarianism, and (what was then) Continental Positivism is quite clear. Still, I thought it might be productive to mash some ideas from Nietzsche into a constructive critique of the Analytic tradition.

Again, this is a very strange sort of conversation I am orchestrating. I am reading Nietzsche, as it were, with apparent innocence. Perhaps I'll be criticized for going outside my sub-sub-specialty, but I figure that when a philosopher publishes a book, even if it's a demanding tome like Beyond Good and Evil, they entitle anyone willing to honestly read it to take a punt at it.

Being creative and understanding creativity: Into the hidden chamber of discovery

Nietzsche is, in a way, sloppy. He contradicts himself a lot. He doesn't have a consistent line on questions like *Is the truth unobtainable, or is it dangerous*? He thinks that the

origins of a thinker's thoughts is a far more telling objection against those thoughts than it really is. Regardless, he has a lot of insights. I am told that he is not the only guy like this in the continental tradition (Fn: Which is not to generalize all continental philosophy as sloppy). Foucault, for example, is sloppy on both history and philosophy, according to philosophers and historians I trust who have read him. Yet, when read in moderation and with a skeptical eye, Foucault offers insights.

In relation to creative endeavors, the argument that sloppiness can be useful isn't controversial. Someone on acid can't reason as well as a sober person, but there may be many forms of creative work for which they are better suited. LSD was instrumental in creating much brilliant music and poetry for example. In such cases, the impairment of rational thought is glued to the rise of creative powers. I see no reason to think this phenomenon is exclusive to the arts. There may well be true or useful characterizations of the world that are easier to find if rational thought is loosened.

Even if you dispute the insights of Nietzsche, he isn't the only possible example. Marxism is a great case study. Analytic Marxism was an attempt to translate Marxism from the continental style into the reasoning style of Analytic Philosophy. It's a cool school of thought, but it's almost impossible to imagine a world in which Marxism was first invented within Analytic Philosophy. Not by accident did Marxism grow on the altogether wilder and more speculative grounds of Left-Hegelian philosophy.

Nor is this phenomenon exclusive to philosophical ideas like Marxism. Many of the great physicists of the early 20th century—think Schrodinger, Einstein, Heisenberg, etc., drew on philosophical ideas that we can politely call "ripe with speculation". It is impossible to say to what extent such speculative ideas were actually *useful* in the scientific thought of these men. Certainly, though, these physicists themselves thought these ideas were useful.

So I suspect the sloppiness and the leaps in logic are active ingredients in Nietzsche's work, allowing him to obtain his biting insights.

If this is right, we face a difficulty. How should we trade off a particular kind of creativity against a particular kind of rigor? How are we to evaluate methodology, if the truth or reasonableness of methodological premises isn't the only factor in evaluation, at least sometimes, and in certain kinds of inquiry?

The best framework I've been able to come up with so far for understanding, though not solving, this impasse is the distinction between *The Contexts of Discovery* and *Contexts of Justification* from the philosophy of science. The context of discovery is the context of creating new good "guesses" about the world (e.g. hypotheses). The context of justification is about justifying—or disproving—those hypotheses.

Analytic philosophy has made great strides in understanding not only deductive reasoning but also inductive and abductive reasoning for and against existing views. Evaluating, understanding, and encouraging the creation of new ideas though, remains largely beyond its reach. In other words, Analytic philosophy has proven much better at understanding the process of justification than it has at understanding the process of discovery—through some are trying to remedy this.

Not only has Analytic philosophy been better at *understanding* justification than at understanding discovery, but it's also arguably been better at *doing* justification (and refutation) than the discovery of new positions. This is most true in relation to social philosophy and theory, and in relation to philosophical psychology. This is by no means to downplay the brilliant new positions that have been created by Analytic Philosophy, enough to fill libraries, but it is something to think about.

By contrast, not only does Nietzsche excel at finding novel hypotheses, he is very much at home theorizing about the context of discovery and ushers us into that place. Nietzsche excels in generating novel and interesting—if perhaps not always persuasive—theories of the genealogies/origins of ideas and institutions. Even more so, reading his own work, we are left with a sense of *philosophy in motion*. Often reading, Beyond Good and Evil, it seemed I was reading thinking, rather than reading the

products of thought—philosophizing, rather than philosophy. There is a kind of teaching about the creative process through demonstration to which he treats us.

So what is the way forward? As I mentioned earlier, evaluating scholarly methodology becomes monstrously harder when it's not a matter of 'just' looking for more truth and rigor in the methodological postulates. I don't have any sense of the right way to balance rigor and soaring when it comes to thinking. My only thought is that one virtue of intellectual diversity is that at least no individual needs to do it all at once.

The human as a believer and the philosophy of belief

I am no historian of thought, but it seems to me that the classical picture of humans in philosophy and "western" thought is that people are *believers*. we believe things, i.e. we have a stock of things that we hold to be true. At least until the modern period, the status of desire in this picture is unclear. Maybe desires are separate things from beliefs, or maybe to desire X is to believe X is excellent or worth possessing. Regardless, belief takes the lead, especially in premodern philosophy which tends to disdain "the passions". Exactly what the passions are—whether they consist in all motivations, all desires, all emotions—what is the relation between these in turn, etc. etc.—is a little unclear. Nonetheless, the passions are certainly distrusted.

Hume comes along and makes the argument that *reason is the slave of the passions*. What he means by this is that no amount of reasoning can ever lead you on its own to want something. In practice, Hume means by this means desire is autonomous from belief. No belief/theory/conception can ever imply a desire/passion/motivation and vice-versa. *[Aside: this is very closely conceptually linked to Hume's further claim that one cannot derive an is from an ought and vice-versa.]*

*S*o we get a bifurcation—there are *beliefs and desires*. In more modern times, formal models of mind and agency have become increasingly sophisticated. There are degrees of belief over different possible states of the world, and utility functions over these same

states. Belief still retains, at least in the context of our thinking about philosophy, a kind of thematic primacy.

I previously have worried a lot about one aspect of this account—the idea that people have clear beliefs. Rather—I think that there are many different components to what we call belief. Often these components come apart—so you can believe in one thing in one sense of belief, and disbelieve it in another sense. Tamar Gendler was one of the first to pioneer this "splitting" of beliefs into multiple components with her concept of Alief, but I think there are many other components as well—at least four by my count.

Nietzsche, I think, turns this picture on its head in a different way. He imagines a philosophy in which the central object is not what we *believe*, but what we *desire* and *will*.

These two critiques—the one I like to push, that belief is fragmentary, and the one Nietzsche likes to push, that our desire is often more central to who we are than our beliefs, complement each other. In many ways, what we desire in our lives and world may be much more stable between contexts, and much less fragmentary, than what we believe about our lives and the world. To the extent that there is any continuity or wholeness to a person whatsoever, it is much more in what they want than in their internal map of the world, still less the stories they tell about why they are doing things.

There is an interesting analogy here, I think, in inverting the usual priority of belief and desire, with the Marxist inversion of Hegel, in which material circumstances and the means of subsistence are seen as having precedence over ideas and ideology. Beliefs stand for ideas, desire for the material basis of life. But that's a topic for another essay. I'll need to do more research because a lot of people have probably already said it.

Jumping ships on moral reasoning

The typical image of moral reasoning in Analytic Philosophy is some variation of the following. Since one cannot derive an ought from an is, one gathers together all of one's

oughts—that is all one's moral urges. Some of these will be contradictory, one seeks to make adjustments to bring them into line with each other, smoothing out real and potential contradictions, and creating a coherent system of moral requirements. This is the method of reflective equilibrium. The philosopher works the jagged clay of moral intuition into a manageable system. I tend to think that even many Analytic Philosophers who claim not to be following this approach covertly are (much to Kieran's dismay).

Although the term reflective equilibrium did not exist at the time, Nietzsche makes it clear that he is unhappy with processes like this. The philosopher shouldn't merely smooth over existing values—they should be a creator of values, or at the very least, an evaluator of values.

But is this possible? Nothing comes from nothing. The only argument in favor of a value can be another value since one cannot derive an ought from an is. The problem, then, with "creating" values without reference to existing values is that we have nowhere to stand.

Otto Neurath, in speaking about beliefs rather than moral principles, said something that can be applied to morals:

"We are like sailors who on the open sea must reconstruct their ship but are never able to start afresh from the bottom. Where a beam is taken away a new one must at once be put there, and for this, the rest of the ship is used as support. In this way, by using the old beams and driftwood the ship can be shaped entirely anew, but only by gradual reconstruction."

But there is, I think, one sense in which we can critically consider our values in a way that goes beyond reflective equilibrium.

Let's say you were to sort through your moral views and come to some rough sense of what you stood for. It's possible that looking out from that vantage, you might decide

that another set of values is *practically dominant*. By this I mean that adopting that set of values, genuinely committing yourself to them, will make you more likely to maximize not only your own current values and goals but also the new values and goals you adopt. The new set of values might be the best option both from the point of view the new values and from your old set of values.

This is a bit abstract, so let me give an example. When I was severely depressed, I found myself contemplating what I stood for. At the time I saw justice as particularly central to my basic, bedrock morality—justice as an important end in itself. I realized though, that this overweening concern for justice was part of what had made me sick and depressed. By leaping from the values I held then, to a system of values based more on compassion, I would be more able to do good both from the point of view of my old, justice-centered values and from the point of view of my new compassion centered values. This practical dominance argument induced me to change. Through a process that was both practical and philosophical, I rejected an old set of values and adopted a new one.

This, I think, is one way in which Nietzsche's ideal of the philosopher as an evaluator and creator of values—and not merely a smoother and summarizer of them—is possible, but without requiring us to step to an archimedean point outside all value.

[Aside: I do not think this situation I have described is especially rare. It seems to me quite common that two systems of values have mostly overlapping goals, but that one system could be better at getting at those goals than the other. Here's another way it could happen. You might find that one system of values is very vulnerable to being rationalized—manipulated to get the result you want—due to the intricacy of its fine distinctions and casuistry. This makes you think you should prefer another code that, in practice, overlaps in its ultimate ends to a large degree with your current code, but is less vulnerable to being rationalized in this sense. Thus, you switch.]

The neglected emotional-aesthetic questions of philosophy

Nietzsche draws our attention to what we might think of as emotional-philosophical problems. Problems that are, to speak somewhat inaccurately, *non-cognitive*.

When I was in high school my teacher, in talking about our essays on King Lear, suggested that we had to grapple with the problem of evil. I was a little perplexed. I was not religious, nor, to my knowledge, was she. What is the problem of evil in a world without God? There exists evil. It's sad. The end.

I still don't know what she meant by posing this, perhaps she didn't mean anything very definite. Having read Nietzsche though, I think I see both the outline of a secular continuation of the problem of evil, and a secular solution.

Nietzsche grapples with Schopenhauer's pessimism. In a very crude sketch, this pessimism goes as follows. Life is a struggle. We seek what we desire. Satisfaction is only a brief respite, leading to the next round of seeking and frustrated desire. The overall picture given by Schopenhauer is very nearly identical to that propounded by the Buddha. To exist as a being is to want things. To want things is to be frustrated and incomplete.

The new, secular, problem of evil then is this, how can we give meaning to this process in a way which makes the continual frenzy of frustrated desire acceptable?

This problem is what I call an *emotional* problem. It's a problem with the aesthetics of life. How are we to understand this overall picture in a way which makes it less horrifying, or, if we cannot do that, how are we to resign ourselves to the horror? It's not a factual question about the way things are. It's not even really a moral question about how we should live. It's a question about how we can keep perceiving life as beautiful and enjoyable.

This kind of aesthetic or emotional problem—of making sense of patterns in a way that allows us to process them or cope with them-, is a lot of what ordinary people mean by philosophy. Sadly, it's a pursuit that's often missing from Analytic Philosophy, because what is being searched for is not so much a hidden truth or even imperative—but a method of setting life to the right kind of narrative.

Nietzsche's solution to the problem of pessimism is that the struggle itself has to be seen as valuable—and not merely success in those struggles. By learning to love the beauty in the struggle to live well, we can reconcile ourselves to the permanent hunger of existence. This is a way of framing the aesthetics of life which makes it something other than ugly, and only cruel. I'm not sure if this is quite *the* solution—but I think it's on the right track.

[Aside: I would add, though I don't know that Nietzsche would agree, that the view that certain kinds of striving against obstacles, might have intrinsic value is no reason to keep brutality, crippling diseases, hunger, etc. Even a life in a utopia of material comfort has forms of striving enough—e.g. for artistic or scholarly greatness, for love, to be a good person, etc. I would rather a world in which people fight for self-actualization in Maslow's hierarchy than struggle for food and shelter. This is really a strong argument against social Darwinism—we must give people space to struggle with the higher and more beautiful problems.

I would add also that none of this is broadly opposed to an account of ethics focused on welfare—so long as welfare is understood broadly.]

This secular problem of evil is just one emotional problem in philosophy, there are assuredly many others. It is right to expect attempts at solutions to such emotional problems from philosophers, and it must be seen as a limitation in the analytic tradition that it has not, hitherto, often attempted to provide them.

In the not too distant future—things which might have been only a bare hint in Nietzsche's day—like the possibility of neurologically eliminating thwarted wanting and feeling altogether, might transpire. In a world in which "wire-heading" is possible—the question of the value of striving and thwarted desire may soon become urgent.

Four parts of belief

A little bit of prodding suggests that beliefs are not so simple as they seem. Consider for example Tamar Gendler's concept of an *Aelief*– a kind of belief-like state. An Aelief, per Wikipedia is:

"…an automatic or habitual belief-like attitude, particularly one that is in tension with a person's explicit beliefs.

For example, a person standing on a transparent balcony may believe that they are safe, but alieve that they are in danger…"

Of course there are other ways of dividing things up. When I was a wee undergraduate philosopher my lecturer gave the following case study. Young Catholic men claim to believe that the sin of self-abuse risks their immortal souls. Yet they engage in it with great enthusiasm. Traditional accounts of belief and rationality have difficulty making sense of this. There are, my lecturer suggested, three possible explanations:

A) They don't really believe that self-abuse will send them to hell (though they believe that they believe this.)

B) They are acting incredibly irrationally.

Or, his preferred option C) Belief is not a single thing. It consists in a complex of behaviours, thoughts and feelings which can be separated out from each other. In this

case, two parts of what belief normally is -acting consistently with a view that X and sincerely asserting that X—come apart.

To speculate a little further, the brain is made up of a bundle of systems which are not as well integrated as we might imagine. Thus it is at least possible that the seeming hypocrisy of these young catholic men arises from different mental systems having different models of the world.

I am reminded of the two streams hypothesis– viz, that there is evidence of strong segregation between the brain pathway involved in visual *awareness* of our environment and the brain pathway involved in visual *action guiding* through our environment.

Here is a list of the separable components of belief I'm aware of. See if you can think of any others:

1. **The non-verbal action component**

If I believe it is raining outside I instinctively grab an umbrella on the way out. If I believe the price of oil will fall tomorrow I won't buy oil now. An important subcategory here is betting behaviour, it may lie somewhere between 1&2.

2. The sincere assertion component

This component of belief is the ability to assert with real sincerity that P is true or that you believe P to be true. As in the case of the young catholic men we described above, it can come apart from the non-verbal action component—especially in matters of sacred belief.

3. The feeling component

Often, someone with an anxiety disorder can assert that something they fear is not true. Their non-verbal actions may also reflect this apparent disbelief, to varying degrees. Yet they are still very distressed by some troubling possibility *as if they believed it were true or going to come true*. We might call this the feeling component of belief. This is primarily what the concept of aeliefs gets at.

4. The commitment component

This one (owing to Kieran Latty) is a bit less well defined than the others. It consists in a determination to believe X, in the sense one or more of the modalities of belief listed above. Suppose for example that I have no real conviction that humanity will survive the next hundred years, but I consider it strategically important that I believe such a thing. I might be committed to making myself act, sincerely speak and feel as if it were true that humanity will survive. Whether or not I succeed in this commitment, I have the commitment component of belief.

A sketch of a layered solution to the interpersonal comparison problem

I can't explain the state that I'm in, the state of my heart, he was my best friend.

-Sufjan Stevens, *Wasps of the Palisades*

My Ph.D. is partly about the interpersonal comparison problem. It's a somewhat nerdy topic, often regarded as insoluble. After explaining it to you, I want to convince you that A) it's a really important question and that B) there is a solution that has been hiding in plain sight. What I also want to do is to write an essay on a technical philosophical problem in a very accessible way. There's a great deal of mystery about what philosophers do, so I want to draw back the curtain—not by description but by example.

The “answer” takes a special form. I first propose a solution. I then say “well suppose you don’t accept this assumption in the solution, if you just accept this alternative, weaker assumption you can still get the result”. Then I weaken it again and so on.

It is my hope to show that, so long as you think that the branch of psychology known as psychometrics is broadly acceptable in its methodology—at least as provisional best practice—and you are willing to make some very weak additional assumptions, the interpersonal comparison problem is a solved one. Or rather, the interpersonal comparison problem is solved *as a practical barrier to ethical inquiry that relies on interpersonal comparisons*. There may still be lingering conceptual questions, but no one should hold these questions up as a reason not to use interpersonal comparisons in ethical inquiry—for example, in thinking about welfare economics.

But I’m getting ahead of myself. Here’s an example to introduce the problem we’re talking about. A hospital is running low on painkillers. There is only enough left for one patient. One has a headache secondary to a head cold. The other has dislocated their shoulder. Which should get the painkiller? Most of us will have zero problems coming to a resolution, the patient with a dislocated shoulder should get it. But what does it mean for one person to be in greater pain than another? How can we quantify pain in a way that can be compared across persons?

Even in the form, I have put it, this is not a purely hypothetical problem. I have worked in reception jobs in hospitals and watched medical staff make decisions about triage etc. partly on the basis of considerations of the degree of pain between individuals. If there’s no scientific way to make those judgments, that’s bad news. If those judgments are meaningless even in principle, that’s even worse.

What I’m speaking of is the problem of interpersonal comparison. It’s not just a philosopher’s question! In economics, at least since Lionel Robbin’s book on the subject, it has been something of an article of faith that interpersonal comparison is difficult, or fraught with difficulties, or perhaps is nothing but the representation of a

decisionmaker's own preferences over the tradeoff rate between people. On the basis of skepticism about interpersonal comparison, welfare economics became all about Pareto improvements and the Kaldor-Hicks criteria (if you don't know what these mean, don't worry, it won't matter here). This, in my opinion, contributed to the perception of economics as an anti-egalitarian science. It may have even led policy in an anti-egalitarian direction.

Practical applications

The problem of interpersonal comparison is a very general one. Consider the field of effective altruism, which aims to maximize the good done with a given set of resources. Good in this field is often conceived of in terms of human (or animal) welfare. Comparing two proposals in terms of their effects on human welfare almost necessarily requires quantifying mental states.

Here's why. Two of the most popular theories of what it means for a person's life to go well are the hedonic theory, according to which one's life goes well to the degree that there is a preponderance of pleasure over pain—and the preference satisfaction theory, according to which one's life goes well to the degree that one gets what one wants. On both these theories of welfare, quantifying the benefits of different welfare improving programs will likely require comparing the intensities of different mental states across different people—pleasures, pains and wants.

There is a third theory of human welfare—the objective list theory—according to which a person's welfare is constituted by the degree to which they have certain good things—like friendship, opportunities, security etc. This might seem to get us out of the problem of having to compare intensities of mental states, but really it doesn't, because, in almost all plausible versions of this account, pleasure and desire satisfaction are important items on that list.

So we can't really do effective altruism without some method—even if only an implicit one—of comparing the intensity of mental states between people.

Beyond effective altruism, consider also the problem of artificial intelligence alignment. Much research is happening at the moment on the question of how to define human ethical priorities formally in such a way that an intelligent machine could be instructed to respect them. On most accounts of ethics, part of our informal, everyday ethical calculus is making these interpersonal comparisons. A clearer understanding of how comparisons can be done in a principled way is thus necessary for AI alignment research.

Defining the problem

Attentive economists and philosophers might have noticed that I speak here of the *interpersonal comparison problem*. I do not speak of the *interpersonal utility comparison problem* which is its more common name. This is for two reasons.

The first is that utility is a poorly defined term. It is sometimes treated as synonymous with *welfare* or *wellbeing* and it is sometimes treated as synonymous with *preference fulfillment* (as in the Von-Neumann Morgenstern utility model). Even more confusingly, these two things—preference fulfillment and wellbeing—are sometimes treated as synonymous with each other and sometimes not.

The second is that there are interesting problems about comparing mental states that may not be directly related to utility at all. I might want to say "Bob is feeling angrier than Alice", and although, of course, Bob's degree of anger is related to both his utility and his welfare, on no definition whatsoever is it constitutive of it. The problem of how to compare Bob and Alice's degree of anger, and the meaning of such comparisons, is an interesting problem in and of itself.

So what we're really interested in comparing is the intensity of certain kinds of mental states between people. Exactly what is in this bundle of mental states is a little difficult to enumerate—but I would put forward, as a basic list:

Pleasures and pains

Desires and aversions

Emotions

There is no need to include *beliefs* at least on a certain definition of belief, as Bayesians have given us an adequate account of how to compare the strength of beliefs using betting behavior.

I call this category *affective mental states* because they all seem to have a tight conceptual link with motivation.

Empirical usefulness and psychometrics

Before certain complexities are added, I don't really think that the interpersonal comparison problem is that difficult. Consider, what makes us think that we can compare temperatures between objects? We develop hypotheses about ways to measure temperatures, and how hot and cold certain things are. We find that using these hypotheses we can do empirical work—make predictions and so on. That's really all it is.

Can a guess about the relative intensity of some affective mental state do empirical work? Can it helps us make true predictions, and not lead us too often to false ones? Yes!

There's a whole science called psychometrics which makes estimates of the magnitude of various mental constructs, including, but not limited to, affective states of all the types we discussed above. Indeed, within psychometrics, there is a field of happiness studies, focused specifically on constructs like life satisfaction and happiness that many consider of one essence with welfare itself.

But these psychometric approaches were historically neglected by economists and philosophers working on the problem of interpersonal comparison. Indeed, psychometric approaches have often been neglected in general in these fields—though this is changing now—see the emerging field of happiness economics and the work of the philosopher Alexandrova.

There's a philosopher called Angner who has been working on the differences between psychologists and economists in the measurement of welfare for a while. His thesis is that it comes down to different understandings in the theory of measurement. Psychologists use a more flexible, one might say, empiricist, approach to measurement called the psychometric approach, whereas economists prefer the representational theory of measurement, a more rationalist approach which is based on formal axiomatizations.

The way psychometrics and the psychometric theory of measurement operates is by assigning magnitudes to a person's level of a construct through tests with standardised items ("Barry's level of happiness is 7/10 whereas Alice's level of happiness is 9/10") *and then using those assignments of numbers to make predictions*.

Let's say that we're measuring happiness. We begin by creating a series of questions that we think, based on our understanding of happiness, should measure happiness.

E.g., *rate the following propositions 1 to 5, with 1 being strongly disagree and 5 being strongly agree.*

1. *I am generally in a positive mood.*
2. *I feel good about life.*
3. *I am a happy person*

We can see that the test has a certain plausibility because its questions are conceptually related to happiness (it has "**face validity**"). Even this alone gives us some basis for credence in the test as a measure of happiness.

Now we administer it to a bank of people, using it to assign estimated happiness scores. We first check to make sure that it is measuring *something* and doesn't just represent random noise, this is to say we check the measure for **reliability.** There are a couple of different ways to do this, but one is to administer the same test to the same group of people with, say, an hour's gap between, and check the correlation between the scores at T1 & T2.

Our next task is to check how well it is performing as a happiness estimator. We might use its estimated happiness scores in a regression model to predict results on other tests which measure similar things (we call this "**convergent validity**"). For example, we might correlate it against a preexisting test of, say, hopefulness. One interesting form of convergent validity is to compare first-person results with third-person results. Have Bob fill out the test, then have Bob's roommate, Alice, fill out the test *as if she were Bob* and see what the correlation is.

Or we might use test results to predict a behavioral outcome like suicide rate or frequency of smiling. This is **criterion validity** of a type we might call **behavioral validity**.

We might also flip things around and see how well circumstances, like an unhappy breakup, can predict our assignment of scores via a test. This would be another example of **criterion validity—**I call this subtype **situational validity**.

We might even develop hypotheses about how our measure should be related to biology if it really does capture happiness. For example, we might check to see if it is inversely related to stress response hormones like cortisol. This would be another type of **criterion validity** we can call **biological validity.**

Thus psychometrics gives us a way to estimate the relative intensity of affective states. It then tests these estimates, seeing if they are borne out in behavior, environment, biology, peer opinion, and other tests. Through an iterating process of testing, theory development, and application psychometrics aim at better and better ways of assigning numbers to mental states in a way that is valid across persons. There's a lot I haven't gotten into here, including more details on the role of statistics—especially factor analysis and psychometrics is not a field without methodological controversy, but, overall, it seems psychometricians never got the memo about the impossibility of interpersonal comparison.

The escape route

To me it seems that psychometrics is measuring *something* interpersonally between people—its capacity for empirical success shows this. Thus, a skeptic of interpersonal comparison owes us an account of what psychometrics is and is not measuring if they are to maintain that interpersonal comparison for ethical purposes is impossible.

The most plausible approach here is to insist that there is a distinction between mental states conceived of in *how we experience them* and conceived of in *how they influence our behavior*. This probably seems very abstract, so let me explain.

Consider the concept of *qualia*. To introduce the idea of qualia consider Alice. Alice has spent her whole life seeing the color spectrum inverted. Her greens are reds. Her yellows are blues.

However, from a young age, she was taught language like everyone else, to associate words with colors she saw. Thus she calls her green experience of what we consider to be a red object "red", just as we do. Presumably, no one will ever even know that Alice's experiences are so very different from ours in this way. This "greenness of green" is what we call qualia, and though it seems immediately present in consciousness, it's hard to imagine what difference it could make to behavior.

It might be a short step from admitting qualia to making interpersonal comparison impossible. Consider the feel of desire, of longing. Now imagine that all your longings and aversions were exactly twice as great. You might think that this would have notable impacts on your behavior—perhaps making you a more passionate person—but there is a strong argument that it wouldn't. For example, your strengthened desire to act might be exactly counterbalanced by your increased laziness. Perhaps then, experience can be altered without a functional alteration in the case of affective states, if you scale them in proportion to each other.

So, in order to prevent the conclusion that psychometrics can be used for interpersonal comparison, what our interlocutor is aiming at is a bifurcation between the functional part of an affective state (which we will call an f-state) and the experiential part of an affective state identifiable with qualia (which we will call an e-state). It is acknowledged that psychometrics can measure and compare f-states, but e-states are more mysterious

and inscrutable—hence thwarting efforts at interpersonal comparison. Remember those terms, e-state, and f-state, they're going to keep coming up.

If you're wondering why psychometrics can only measure f-states, remember that what psychometrics measures is behavior (even if it's only question answering behavior), and that if it influences behavior, it's part of the f-state.

There's a further assumption here. The critic assumes that it's these scientifically inscrutable, interpersonally incomparable e-states that matter for ethical purposes—it's these states which comprise human welfare or suffering. If we acknowledged e-states existed but didn't regard them as ethically important, they wouldn't be troubling from the standpoint of ethical decision-making or policymaking. Thus while they would create difficulties for interpersonal comparison, they wouldn't be difficulties of practical relevance.

Cutting the escape route off at the pass: functionalism

In the previous section, I explained a way out of the seemingly obvious conclusion that psychometrics enables interpersonal comparison. That escape route was to disentangle feeling and behavior in a particular way. In this section, I'm going to outline a counterargument against this "escape route".

Functionalism is a view in the philosophy of mind about what the mind is. It might be best to explain it by way of comparison to analytical behaviorism because it can be seen as a more evolved version of that doctrine.

Analytic behaviorism, a now almost extinct view in the philosophy of mind, held as follows. Let's say you are angry. That anger is *constituted* by certain behaviors and behavioral tendencies. For example, you may raise your volume, tend to act

destructively and rashly, become flushed in the face, etc. Those behaviors and behavioral tendencies are your anger. Analytical behaviorism has the advantage of being a purely physical view of what the mind is, but it has disadvantages. For example, we generally think that your anger *causes you* to raise your voice. But if your anger is *partially constituted* by your tendency to raise your voice, it's not really accurate to say that your anger has *caused you to raise your voice*.

The functionalist has a solution to these and many other problems of analytical behaviorism. What if your anger is *whatever arrangement of your central nervous system that causes you to behave in an angry way*? This keeps a tight conceptual connection between behavior and mental states while making sense of our ordinary intuition that mental states cause behavior.

Functionalism abolishes the possibility of a residual unobservable difference in mental states by holding that e-states separate from f-states don't exist.

There are many good arguments for this kind of functionalism that denies there are separate f-states and e-states. Consider, for example, that if e-states truly are separate from f-states, they have no influence on behavior since f-states can include anything that has an effect on behavior. The theory then faces a problem ***why are we talking about e-states if they have no influence on behavior?*** (For those interested, this objection mirrors a classical objection to epiphenomenalism in the philosophy of mind).

So if you accept functionalism, your confidence in psychometrics as a yardstick of interpersonal comparison will once again be restored.

Epistemic functionalism

But okay, okay, I'll admit, not everyone is going to be persuaded by my hardline view that all affective states are functional and contain no non-functional components, but I can sweeten the pot, or rather, remove a lot of the vinegar.

We can weaken functionalism considerably from a claim about how things are to a claim about what it is reasonable to believe (an epistemic claim). If functionalism is the principle that *it is a metaphysical truth that no functional differences=no mental differences*, epistemic functionalism is the view that *it is at least reasonable to assume that there are no mental differences where there are no functional differences unless shown otherwise.* Epistemic functionalism is a weaker claim, functionalism implies epistemic functionalism but not vice-versa. By making our premises weaker while still trying to reach the same conclusion (a common strategy in philosophy), we're trying to make an argument that's appealing to a broader circle.

To further explain epistemic functionalism, let's go back to the example of color experiences (even though it's not strictly related to the problem we're considering here). It could be that you see green where I see red and vice versa, but until someone comes up with evidence of that, it's not irrational to think that your green is much like mine and vice versa. connecting it to our topic, perhaps it is possible that all your emotions or all your desires are on a different scale to mine, but epistemic functionalism suggests that we can reasonably assume they are similar in the absence of contradictory evidence.

I'm going to label the rest of this essay as an appendix because it gets more complicated from here on out, and I think that for many people the arguments I have made thus far will go through. Nonetheless, keep reading if you want to learn how we can weaken the assumptions we've made even further.

APPENDIX: TWO EXTRA ARGUMENTS

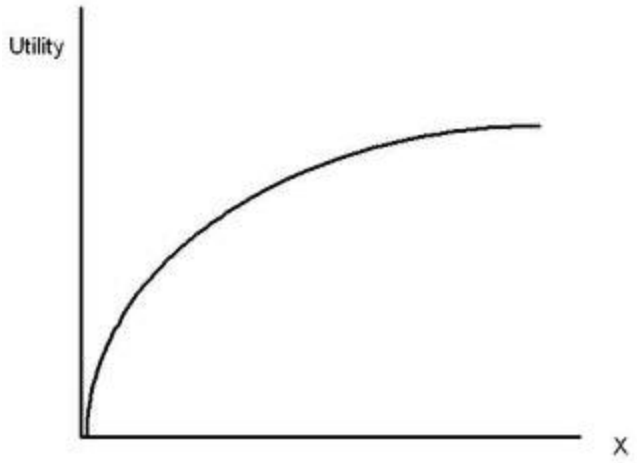
If even still you don't accept this?: Unbiased estimator functionalism

I find the argument so far persuasive as a solution to the interpersonal utility comparison problem. I'm a functionalist. I think mental states are definable in terms of functional relationships with behavior, and hence are fully psychometrically measurable. Even if I weren't a functionalist, I would find epistemic functionalism or the view that it's reasonable to assume that similar f-states equals similar e-states in the absence of contradictory evidence persuasive.

However, I think we can add another layer of "even if". Even if you find all of the above reasoning unpersuasive, an old argument called the equal ignorance argument, combined with an even weaker form of epistemic functionalism that I call unbiased estimator functionalism, might still go through.

Unbiased estimator functionalism: The equal ignorance principle

Let's suppose that earning another dollar always makes you better off—this means that your utility function is **strictly increasing** in dollars. Let's further suppose that, despite this, each additional dollar is worth less to you than each previous dollar—this means that your utility function is **concave** in dollars. Perhaps your interest in dollars looks something like this:



Now let's suppose that a decision-maker knows that every single person in the population has these features—utility which is both strictly increasing and concave in dollars. However, the decisionmaker has no further information on the utility functions of the population—their shape or magnitude. The decision-maker has a pool of money that they want to dole out to the population, how should they divide it?

You may be able to see the answer—which Abba Lerner proved mathematically—intuitively. The decision-maker should split the money equally.

But the equal ignorance theorem suffers from a problem if it is meant to apply to real life. Arguably we are not totally ignorant about the scale of other people's utility functions. For example, arguably we have good reason to think that, on average, rich people like money more than poor people *because all other things being equal, the person who likes X more will have more of it*. We might also have other reasons to think there are differences. For example, the rich might be more habituated to their wealth—this could create in them a greater need for it or might mean that it is largely wasted on them.

Unbiased estimator functionalism: The claim

By an unbiased estimator here we mean something slightly different from its normal usage. We mean an estimator that, in expectation, is not biased towards higher or lower values for any person or type of person. We still may not trust the answers it gives, we may think it tends to be wildly inaccurate—but not in the form of a known bias.

Let's say you have two friends named Alice and Bob. If you tell either of those friends a story about someone, they will then guess that person's height (it's very irritating). They are both very bad at it and are on average wrong by two feet. However, when Bob makes predictions, he tends to overestimate the height of men, and underestimate the height of women. Alice has no specific tendency towards underestimating or overestimating the heights of men or women. She's just all over the shop. Alice is an *unbiased estimator*.

Here's where unbiased estimator functionalism comes in. According to unbiased estimator functionalism, a person's functional state can be used as an unbiased estimator of the intensity of their affective states. F-state is an unbiased estimator of e-state.

Glossing quite a bit, the argument from this to interpersonal comparison goes as follows. Let us suppose that we think f-states are measurable and interpersonally comparable, but that e-states are what matters in ethical terms.

If we accept that f-states are unbiased estimators of e-states, even if we don't think they're necessarily very accurate estimators, and that we have no further information about the relative magnitudes of e-states, then, using reasoning like that involved in the equal ignorance argument, we can derive the conclusion that our estimates of e-states based on f-states should guide our ethical behavior. We may be wrong, but our

wrongness has no tendency to go in a specific direction and we have no further information so we can't do better than just using f-states as an estimate of e-states.

One final even if

Maybe you've read through all the above and you find yourself radically uncertain about whether or not the arguments I have given "go through" in preserving interpersonal comparisons. I think the argument has been compelling, but I get stuff wrong all the time. Well, I have one final pitch for you, a pitch to commonsense.

Denying interpersonal comparisons is a kind of skepticism. There are all sorts of arguments for all sorts of forms of skepticism in philosophy. For example, some people think that we should be skeptics about induction—or making inferences about the future on the basis of past observations. These people point out that the principle of induction—that it is legitimate to make such inferences—is itself undefended *and it's not good enough to say that it must be true because it's worked in the past, because that's circular!*

Generally speaking, when we run into a theoretical argument for skepticism in philosophy, we don't allow it to stop actual everyday and scientific inquiry. We assume the philosophers will work it out someday, and even if they don't, we likely keep going all the same.

We have a practically applicable method for making interpersonal comparisons of affective states (psychometrics) that lines up with commonsense ideas about how to measure a variable between instances, and how to validate that measure. Even if you believe that there are philosophical reasons to be skeptical of psychometrics as a strategy for interpersonal comparison, I propose we should keep using it in the interim,

much as we keep assuming the future will resemble the past, other people have minds, there is an external world, and so on.

Recent advances in Natural Language Processing—Some Woolly speculations

I wrote this essay back in 2019- before GPT-3. Since then I think it has held up very well. I thought I'd re-share it to see what people think has changed since then, in relation to the topics covered in this essay, and see if time has uncovered any new flaws in my reasoning.

Natural Language Processing (NLP) per Wikipedia:
"Is a sub-field of linguistics, computer science, information engineering, and artificial intelligence concerned with the interactions between computers and human (natural) languages, in particular how to program computers to process and analyze large amounts of natural language data."
The field has seen tremendous advances during the recent explosion of progress in machine learning techniques.
Here are some of its more impressive recent achievements:

A) The Winograd Schema is a test of common sense reasoning—easy for humans, but historically almost impossible for computers—which requires the test taker to indicate which noun an ambiguous pronoun stands for. The correct answer hinges on a single word, which is different between two separate versions of the question. For example:

The city councilmen refused the demonstrators a permit because they feared violence.
The city councilmen refused the demonstrators a permit because they advocated violence.
Who does the pronoun "They" refer to in each of the instances?

The Winograd schema test was originally intended to be a more rigorous replacement for the Turing test, because it seems to require deep knowledge of how things fit together in the world, and the ability to reason about that knowledge in a linguistic context. Recent advances in NLP have allowed computers to achieve near human scores:(https://gluebenchmark.com/leaderboard/).

B) The New York Regent's science exam is a test requiring both scientific knowledge and reasoning skills, covering an extremely broad range of topics. Some of the questions include:

1.Which equipment will best separate a mixture of iron filings and black pepper? (1) magnet (2) filter paper (3) triplebeam balance (4) voltmeter

2. Which form of energy is produced when a rubber band vibrates? (1) chemical (2) light (3) electrical (4) sound

3. Because copper is a metal, it is (1) liquid at room temperature (2) nonreactive with other substances (3) a poor conductor of electricity (4) a good conductor of heat

4. Which process in an apple tree primarily results from cell division? (1) growth (2) photosynthesis (3) gas exchange (4) waste removal

On the 8th grade, non-diagram based questions of the test, a program was recently able to score 90%. (https://arxiv.org/pdf/1909.01958.pdf)

C)

It's not just about answer selection either. Progress in text generation has been impressive. See, for example, some of the text samples created by Megatron: https://arxiv.org/pdf/1909.08053.pdf

2.

Much of this progress has been rapid. Big progress on the Winograd schema, for example, still looked like it might be decades away back in (from memory) much of 2018. The computer science is advancing very fast, but it's not clear our concepts have kept up.

I found this relatively sudden progress in NLP surprising. In my head—and maybe this was naive—I had thought that, in order to attempt these sorts of tasks with any facility, it wouldn't be sufficient to simply feed a computer lots of text. Instead, any "proper" attempt to understand language would have to integrate different modalities of experience and understanding, like visual and auditory, in order to build up a full picture of how things relate to each other in the world. Only on the basis of this extra-linguistic grounding could it deal flexibly with problems involving rich meanings—we might call this the multi-modality thesis. Whether the multi-modality thesis is true for some kinds of problems or not, it's certainly true for far fewer problems than I, and many others, had suspected.

I think science-fictiony speculations generally backed me up on this (false) hunch. Most people imagined that this kind of high-level language "understanding" would be the capstone of AI research, the thing that comes after the program already has a sophisticated extra-linguistic model of the world. This sort of just seemed obvious—a great example of how assumptions you didn't even know you were making can ruin attempts to predict the future.

In hindsight it makes a certain sense that reams and reams of text alone can be used to build the capabilities needed to answer questions like these. A lot of people remind us that these programs are really just statistical analyses of the co-occurence of words, however complex and glorified. However we should not forget that **the statistical relationships between words in a language are isomorphic to the relations between things in the world—that isomorphism is why language works**. This is to say the patterns in language use mirror the patterns of how things are(1). Models are transitive—if x models y, and y models z, then x models z. The upshot of these facts are that if you have a really good statistical model of how words relate to each other, that model is also implicitly a model of the world, and so we shouldn't surprised that such a model grants a kind of "understanding" about how the world works.

It might be instructive to think about what it would take to create a program which has a model of eighth grade science sufficient to understand and answer questions about hundreds of different things like “growth is driven by cell division”, and “What can magnets be used for” that wasn’t NLP led. It would be a nightmare of many different (probably handcrafted) models. Speaking somewhat loosely, language allows for intellectual capacities to be greatly compressed *that's why it works*. From this point of view, it shouldn’t be surprising that some of the first signs of really broad capacity—common sense reasoning, wide ranging problem solving etc., have been found in language based programs—words and their relationships are just a vastly more efficient way of representing knowledge than the alternatives.

So I find myself wondering if language is not the crown of general intelligence, but a potential shortcut to it.

3.

A couple of weeks ago I finished this essay, read through it, and decided it was not good enough to publish. The point about language being isomorphic to the world, and that therefore any sufficiently good model of language *is* a model of the world, is important, but it’s kind of abstract, and far from original.

Then today I read this report by Scott Alexander of having trained GPT-2 (a language program) to play chess. I realised this was the perfect example. GPT-2 has no (visual) understanding of things like the arrangement of a chess board. But if you feed it enough sequences of alphanumerically encoded games—1.Kt-f3, d5 and so on—it begins to understand patterns in these strings of characters which are isomorphic to chess itself. Thus, for all intents and purposes, it develops a model of the rules and strategy of chess in terms of the statistical relations between linguistic objects like "d5", "Kt" and so on. In this particular case, the relationship is quite strict and invariant- the "rules" of chess become the "grammar" of chess notation.

Exactly how strong this approach is—whether GPT-2 is capable of some limited analysis, or can only overfit openings—remains to be seen. We might have a better idea as it is optimized — for example, once it is fed board states instead of sequences of moves. Either way though, it illustrates the point about isomorphism.

Of course everyday language stands in a woollier relation to sheep, pine cones, desire and quarks than the formal language of chess moves stands in relation to chess moves, and the patterns are far more complex. Modality, uncertainty, vagueness and other complexities enter- not to mention people asserting false sentences all the time- but the isomorphism between world and language is there, even if inexact.

Postscript—The Chinese Room Argument

After similar arguments are made, someone usually mentions the Chinese room thought experiment. There are, I think, two useful things to say about it:

A) The thought experiment is an argument about understanding in itself, separate from capacity to handle tasks, a difficult thing to quantify or understand. It's unclear that there is a practical upshot for what *AI can actually do*.

B) A lot of the power of the thought experiment hinges on the fact that the room solves questions using a lookup table, this stacks the deck. Perhaps we be more willing to say that the room as a whole understood language if it formed an (implicit) model of how things are, and of the current context, and used those models to answer questions? Even if this doesn't deal with all the intuition that the room cannot understand Chinese, I think it takes a bite from it (Frank Jackson, I believe, has made this argument).

— —

(1)—Strictly of course only the patterns in true sentences mirror, or are isomorphic to, the arrangement of the world, but most sentences people utter are at least approximately true.

The Paradox of the Crowd

Consider the tension between these statements:

1. The majority opinion in almost every field is more likely to be correct than your own, if your opinion deviates from the majority opinion. This is true even if the group has no more raw data than you—because the aggregate reasoning of the group is likely to be better than yours.
2. If, on the basis of (1), everyone reasoned that they should simply adopt the majority view, the quality of the majority opinion would fall.

I take it that it is obvious that the above statements are usually true. (1) may not always hold—for example you might be massively better informed and better at reasoning about a topic than everyone else with an opinion on it. But in the main, individuals are not more likely to be correct than the majority opinion. Even if you think you are an exception because you are better informed and cleverer, which is more likely- that you truly are an exception, or that you are one of the many people deluded into thinking you are an exception?

Admittedly, this is all very context dependent. If you're the only biologist in your society and everyone else has no scientific training and is a creationist, you probably shouldn't be worried about the above paradox, because you'll simply reject (1). The case we have in mind is instead that of an inquirer in a community of relative epistemic equals.

To dramatise the paradox: Galileo said that "In questions of science, the authority of a thousand is not worth the humble reasoning of a single individual." This is wrong, the humble reasoning of a single individual is usually not that good. What is true is that without the humble reasoning of individuals and small groups, the authority of a thousand would not advance.

One way forward here is to create two sets of propositions. The set of propositions we 'believe' in the sense that we conduct our investigations on the basis of them, and use as the basis of our arguments internally to a community of inquiry and a second set of propositions we 'believe' in the sense that, if you were asked what was all things considered most likely, you would assent to. We populate the first category with propositions that *seem to us to be true on the basis of all available evidence except the evidence of the judgments of others*. We populate the second category with propositions that *seem true on the basis of all available evidence including the judgements of others*.

Of course belief partitioning in this way may not be psychologically viable for individuals. One alternative would be to give individuals tacit permission to engage in self-deception about the likelihood that they've grasped something the majority hasn't.

Then there are hybrid models, where we put some weight on the reasoning of others, but not as much as it probably deserves. This would have the effect of preventing too much ink and lucre being spent on fringe ideas, while still alleviating the paradox of crowd judgement.

Part of orienting ourselves in this landscape is to reflect on our goals. *Are we trying to be right, or are we trying to make the group that we are a part of right?* Traditional epistemology has assumed that the goal of the agent is, or should be, to have correct beliefs about the world. To this end they seek to form justified beliefs. What if instead we view our goal as trying to expand the knowledge of the group as a whole? This can be quite a liberating way of seeing things. Got some eccentric hobby horse ideas? Excellent! Someone needs to follow those up. Act like you don't know how much of a stab in the dark it is, or even fool yourself into believing they're likely true if it helps. It's all to the good and expands the cognitive reach of the group.

We now enter the domain of *Normative Social Epistemology, the* study of reasoning for, and as a part of, a group of enquirers, with the aim of supporting that group in its collective search for truth.

Why I left philosophy

1.

I started working on intuitions. To see what a philosophical intuition is (or rather, what one type of philosophical intuition is), consider the following:

You might think knowledge is justified and true belief. But suppose I look at my watch and it says the time is 12:37. On this surely reasonable and justified basis I believe that the time is 12:37, and indeed the time is 12:37. However, unbeknownst to me my clock is stopped. It just so happened to stop on 12:37, and by coincidence this happens to be the time now.

Many people have the intuition that in such a case you do not know that the time is 12:37, but you are justified in believing it, your belief is true, and you certainly do believe it. Thus, they argue, having a justified true belief does not guarantee knowledge. If this is true, it overturns what was the almost universally accepted view of what knowledge almost two and a half millennia—that knowledge is justified true belief, often shortened to JTB. This sense of wrongness about the idea that the person in the example knows that it is 12:37 is a paradigm case—perhaps the defining example—of a *philosophical intuition*. A *philosophical intuition* is typically (and these are neither necessary nor sufficient conditions!) a sense of rightness or wrongness about the application of a predicate—for example "Knowledge" in a hypothetical case. This sense of rightness or wrongness does not seem to rely on anything external to itself for its own justification, rather it just sort of seems self-evident.

But why are we confident in our intuitions? Why should these sorts of intuitions count as evidence? Do we all share the same intuitions about various philosophical topics?

During my undergraduate years, debate had erupted in the literature about what is sometimes called experimental philosophy. Experimental philosophers run experiments that seem to suggest a diversity of sometimes contradictory intuitions between cultures, social groups, and even within individuals about the same cases. For example, there is some evidence that East-Asian subjects actually regard Gettier type cases as instances of knowledge—i.e., they would grant that the person whose watch says 12:37 knows that it is 12:37, even if this only happens to be the correct time by accident. I wrote a senior thesis arguing against two views on what these experiments meant. On one hand I argued against people who say that these results debunk the method of hypothetical cases philosophers' use. On the other hand, I argued against philosophers who contend that the intuitions of non-experts about these matters should not be trusted.

My argument was informed by an underlying thesis about what intuitions were. I regarded intuitions not as glimpses into a platonic realm, but as constitutive and at least partially stipulative. It was my view that intuitions do not *track* the truth about philosophical questions, but rather that they are partly responsible for *making the truth* about philosophical questions.

To see how, let us come back to the Gettier case—although what I say here could apply to many other debates in philosophy. If Janet has an intuition that the Gettier case is not knowledge, this expresses Janet's predisposition to define knowledge in such a way so as to exclude these cases—that she refuses to use the word "knowledge" in a Gettier case partially constitutes what she means by "knowledge".

If Jiang has a a conflicting intuition, Jiang is demonstrating that she is attaching a different *concept* to the word "Knowledge". To talk about Janet being right or Jiang being wrong would be nonsense, because their intuitions pick out different concepts. It would be a bit like an argument about whether it was raining on Wednesday where one person means last Wednesday and the other person means this Wednesday. Jiang and Janet could argue about or investigate whose version of the concept was closer to the typical version, and what sort of people use what variants from the big family of Knowledge concepts—and to do so they might find experimental philosophy useful. They might even debate which concept of knowledge was more *useful* but debating which is correct is meaningless. Each of the two concepts describes or picks out a different property. The role of experimental philosophy then is to show us the variations in the concepts people are deploying—it is a project of philosophical lexical semantics.

A lot of people seemed worried that this approach would lead to anything goes relativism. I don't see the argument. If Jiang has X intuitions about the concept of knowledge, and Janet has Y intuitions about the concept of knowledge, then they are using slightly different words. There will still be a fact of the matter about whether someone's belief is knowledge as Jiang means it, and there will be a matter of fact about whether someone's belief is knowledge as Janet means it. Once propositions are properly disambiguated, there's no spooky 'the world is just a point of view' relativism going on.

There's a common cousin of the view I've outlined that I'd best explain so as to distinguish it from my own. It goes like this. "Philosophers can't study knowledge itself using intuitions, they can only study people's ideas of knowledge. There is this further thing which is Knowledge-itself and either philosophers can't study it, or they need to use some method other than intuitions about cases to do so". I can see why people would confuse this view with the view I've outlined, but I think they're quite distinct. Here's why: If there is a rich enough infinitude of properties in the world, then for any meaningful concept there will be a property corresponding to that concept. As a result, if we come to fully map out someone's intuitions which define a concept C, then we will also find out exactly which property C picks out. If we fully map out a folk concept, then, at least prima-facie, we also fully understand it's corresponding property in the world. There's no need to say things like 'we understand what this group of people mean by

knowledge, but not what knowledge itself is—there's some further fact about that.' What the property of knowledge is, is given fully by what the concept of knowledge is. The character of the property of 'being knowledge' just falls right out of an analysis of the concept of knowledge.

The view I've outlined has many advantages. It avoids tricky epistemological puzzles, for example: why should we think intuitions tell us something about an intangible metaphysical world external to our own minds? If intuitions really do systematically vary between cultures, it avoids the awkwardness of having to explain how one group came to be right and the other group wrong. Also, when you think about it, the idea of one property out there in the world being Knowledge with a capital K is kind of silly. There are of course also arguments against it—mostly in a bundle of ideas related to reference magnetism, direct reference and the causal theory of reference which I won't get into here. There are also concerns built on Quine inspired semantic eliminativism.

Disclaimer for experts: So it doesn't seem like I'm denying certain facts let us clearly acknowledge that nothing in this view says that some linguistic frameworks aren't better equipped to describe the world, or carve it at its joints, or simply be more useful, than other frameworks. Let us also acknowledge that what is 'useful' will be context dependent. Let us acknowledge even further that nothing in our view denies that there may be a degree of reference magnetism towards the 'joints' of nature, it only requires that it not be strong enough to outweigh the possibility of alternative or parallel concepts for important philosophical topics like 'knowledge', 'mind' and 'personhood'.

As far as I know there is no name in the literature for the view about what intuitions are, and what studying concepts really does, that I have described. Despite that, when you talk to philosophers it becomes clear that the view I've described is extremely common. Indeed in some groups, something like it seems to be the majority view, which makes the absence of an explicit name for the view all the more mysterious. The view described here some has similarities with ordinary language philosophy, though few ordinary language philosophers stated it so baldly, and it's hard to tell since, following Wittgenstein, so many ordinary language philosophers seem to have made it a virtue to not be clear on exactly what it was they were doing. It has some definite similarities with what is sometimes called the Canberra Plan. Alvin Goldman outlines a vaguely similar view, although from memory his view is a bit more like the common cousin I outlined a few paragraphs above.

Nonetheless the view is rarely argued for in explicit terms. To revise after so many years I reread the Stanford Encyclopedia of philosophy article on intuitions before writing this—nothing like the view that intuitions make, rather than track, philosophical truths

is described, despite exhaustive discussion of the minutiae of various other views about, and aspects of, intuitions.

As a step towards overcoming this marginalisation, let us call the view *constitutivism about intuitions*. Formally stated then, constitutivism is the view that intuitions are not merely truth-trackers, but truth-makers regarding certain philosophical claims. The reason for this truth-maker status is that a person's intuitions help constitute what they mean by a term in their idiolect, and philosophical questions on the constitutivist view often turn on semantics. Further, from the content of concepts we know trivially the nature of the properties they pick out, for any meaningful, complete and non-self contradictory concept has a property which corresponds to it.

Succinctly—intuitions are part of concepts and different intuitions about appropriate usage give different concepts, concepts give us properties, thus once we've finished conceptual analysis, we've finished. Variation in intuitions , whether between subjects, or even within a single individual, just indicates a plurality of concepts.

2.

But what does this have to do with me leaving philosophy?

I finished up my senior thesis which briefly touched on these themes, and then began a doctoral thesis. In the part of the world I was doing my doctorate, you immediately begin researching and ideally writing your thesis almost as soon as you enrol, unlike the American system. Like any prototypical twenty-two year old philosopher I was wildly over confident in the value of what I had to say, and ready to change the world.

That's when I encountered the haze. A lack of concrete research on the topic of my thesis I could tap into, of an accessible bedrock of literature which I could build a thesis on. There were many papers on metaphilosophy tangentially related of course, but everyone seemed to be coming at it from different angles, groups of people were having conversations that slid entirely past each other. There was no obvious way for me to slip into the party with grace. It seemed to me that there were a great many people who thought they were talking about the same things, but really were talking past each other. There weren't even always names for the various constellations of positions people took. I was lost.

There are, I think a number of causes for the haze. Two of which lay with me, one of which I could be rightly blamed for, and the other I couldn't. The one I could be rightly

blamed for was that I was nowhere near as fucking smart as I thought I was. The one I can't be blamed for was that I was severely mentally ill. However there were also external cases for the haze.

First of all, many philosophers just don't care that much about metaphilosophy. When I talked about my thesis with faculty members and fellow students I generally got one of four responses:

"This is obviously true."

"This obviously not true."

"This is too abstract to be interesting."

"That's really interesting, I'd never thought about it like that."

I realised with astonishment that many—though not all—philosophers had the same orientation towards metaphilosophy that many scientists have towards philosophy. Just like a room of scientists asked about science would tend to throw up sentiments like:

"That's all good and well, but what's the practical point?"

"I have strong opinions on this topic but haven't thought about it that much."

"Ooh that's really interesting but it's a little abstruse for my taste."

So to did philosophers. The other cause of the haze was that even though what I was saying corresponded to what a lot of philosophers believed *no one had thought to name it*. And for some reason, even though many philosophers are instinctively drawn to the idea, people who work on topics like 'what are intuitions', don't seem to much like the idea.

I began to wonder if maybe it wasn't because there's vaguely self-effacing about the whole thing. If you sort of suspect that parts of philosophy (not all, mind) are continuous with semantics—"semantics" being a byword for irrelevancy in our society—maybe you sort of want to shut up about it. Thus you'll only write on relevant topics in meta-philosophy if you think you've got something to say which carves philosophy into a more imposing figure.

I guess I never saw it that way, because I think there's a kind of glory to be had in studying such important concepts and words as *knowledge, personhood, belief, desire* and so on. Some questions about the semantics of some words and the structure of some concepts go to the very heart of the way humans understand their world, and are humans not noble in reason and infinite in faculty?

Overwhelmed by a conversation that was at once fragmentary, vast, and hard to find, in which people who agreed with me never seemed to speak up, which many philosophers seemed to think was a useless conversation, and sapped of energy by my health problems, I sank out of academia.

Against Libertarian Criticisms of Redistribution

Pt 1: Non Aggression tells us nothing about the morality of redistribution

According to the non-aggression principle one should never interfere with the person or legitimate property of another without their permission, unless they have initiated aggression against one first. The non-aggression principle is sometimes taken to be a master argument for libertarian views against the redistribution of money or property—e.g., left wing proposals to redistribute money from the rich to the poor. I won't argue either for or against the principle of nonaggression, as there are far more pressing ethical issues. Instead I'll be contending that the non-aggression principle tells us *nothing*, at least directly, about the topic of redistribution.

In the definition of the non-aggression principle I insisted that the non-aggression principle applies to *legitimate* property. I'm not trying to smuggle anything especially controversial in here. By insisting on the term *legitimate* I'm simply insisting that you actually have to rightfully own the thing in question, it's not enough to simply proclaim that one owns it. A moment's reflection will show that this stipulation is necessary, if one owned everything one proclaimed one owned then many things would have multiple inconsistent ownership claims.

Consider the case of Bob. Bob passionately claims that he owns the Atlantic ocean, he actually seems to believe this, and insists that no one should cross the Atlantic without his permission. When asked to justify this, he responds by saying that crossing his ocean without his permission is aggression, and everyone should accept an ethical norm against aggression. When confronted with this argument, there is no need to say anything for or against the non-aggression principle, one simply has to say that the Atlantic Ocean is not actually Bob's, therefore no aggression against Bob has occurred.

This is where the champion of the non-aggression principle as a basis for libertarianism hits a problem. The supporter of redistributive taxation typically does not accept that the goods and monies to be redistributed are, in fact, the legitimate property of those they are being taken from. They hold, on the basis of a differing theory of distributive justice than that held by the libertarian, that they are the rightful property of someone else.

The libertarian will respond by insisting that, yes, the prior owner is the legitimate owner of the goods or monies in question, but notice that the argument has now strayed beyond the issue of non-aggression into a debate about who owns what. Our point is simple then, non-aggression tells us nothing about redistribution unless we assume that redistribution is a process of removing something from its rightful owner and giving it to someone else *but this is part of what is under dispute in debates about distributive justice*. The debate is really about who is the rightful owner of what, and unless one can win this debate, one might as well be Bob insisting that he owns the Atlantic. Just as there is no aggression against Bob implicit in sailing across the Atlantic ocean and 'breaching' his sovereignty over that ocean, so perhaps there is no aggression in 'taking' money off Beezos to pay for redistribution, if the recipients of that redistribution are already the rightful owners of that money.

Put simply, taking your stuff is not aggression unless it actually does rightfully belong to you, and the whole project of the advocate for redistribution is to try and prove that, in some cases, it doesn't.

In fact if the supporter of redistribution is correct about who rightfully owns what, then in the non-aggression principle would imply that action resisting redistribution is impermissible, as it would be a form of aggression.

Now of course the libertarian has responses to the advocate for redistribution. They can critique the arguments in favour of redistribution and propound their own theories of who owns what that do not allow much of a role for redistribution, for example, as Nozick does in Anarchy State and Utopia. However such arguments are not primarily appeals to non-aggression, rather they are theories of who owns what. Non-aggression simply doesn't cut at the difference between the libertarian and the redistributionist.

Pt II: History and Property Rights

In the previous part I outlined why I don't think non-aggression is sufficient grounds to prove that redistribution is bad. I said that what the libertarian really needs is a theory

of entitlement to property that provides a defence against redistribution. Of course many libertarians have aimed to produce such a theory.

Here are two desiderata for a theory of property rights that can act as the basis of a case against redistribution:

D1. It must justify the existing distribution of property.

D2. It must do so without appealing to the state or any collective body for legitimisation.

Let's unpack desideratum one (D1) a little. It is not enough to establish that people have property rights or that property rights are important to fulfil D1. One can imagine proving that there are lots of property rights, but that they do not correspond to the existing socially recognized distribution. Such a conception of property rights would not be a defence against redistribution, instead it would *require redistribution*. Instead the opponent of redistribution must simultaneously prove that property rights exist, and that they are similar to the existing set of socially and legally recognized property rights.

Desideratum two (D2) is equally important. A proof that property rights exist because the state makes it so would presumably leave the state in a position to change the distribution of these property rights, thus it would be no defence against redistribution.

The most common libertarian approaches which aim to meet these constraints are historical theories of distributive justice. These theories typically hold that you are entitled to something if you justly acquired it from nature, or if you acquired it consensually from someone who did acquire it justly from nature, or if you consensually acquired it from someone who acquired it consensually from someone who justly acquired it from nature, and so on.

I'm sceptical of the claim that any historical theory of distributive justice will ever meet both desiderata because existing regimes of property rights have been arrived at through morally contorted historical processes that libertarians do not accept the validity of. There is no just chain of transmission for the computer I am typing this on. It was built using raw materials from land that was many times stolen and re-stolen. The company that made it was funded using government subsidies that libertarians object to. Every good was made using numerous other goods, and all of those goods made with many other goods in turn, and you don't have to go back far in the history of anything to find numerous interferences with what libertarians would regard as just.

These problems ripple through the whole. If I purchase something from you that you don't really own, I don't really own it either, and if I then sell that thing in turn, I don't really own the money I gain from selling it and so on. Call this the problem of spreading contamination.

At this point most libertarians I've spoken to have responded with something along the lines of the following:

'look, it's the best we've got—sure actually existing society is riddled with theft and misappropriation, and this can be found in the history of practically any consumer goods, land or capital that is owned by anyone, we have to go with what we've got because the alternative is even more theft and misappropriation.'

The problem here is that by definition it's not actually theft unless you own it, and on the historical theory of distributive justice you don't own it because it wasn't acquired justly. You might say "so you're alleging that taking anything from anyone isn't theft—that's absurd." My reply is that I'm saying no such thing, what I am saying is that, taking the historical theory of distributive justice seriously, this is what it entails. This is not an argument that you can take anything you like from anyone; it's an argument that we need a better theory than the historical theory of distributive justice to explain why you can't.

Appendix A: The contention that objecting to capitalism is objecting to private arrangements purely between private individuals

Sometimes Libertarians argue that capitalism is just a series of private arrangements between consenting adults, so there is no grounds to dissent from it. The thing is, most actual exchanges under capitalism involve claims to capital goods and land *that society might well contest the ownership of.*

Let's suppose I found a television by the side of the road, dumped there by some thieves, and was about to sell it. Whereupon you discovered me, and explained that, actually, you have a better claim to the television, and so the transaction can't go ahead.

If I then reply "but this is a contract between consenting adults!" this would be wholly irrelevant *because our agreement involves infringing the rights of others*. The libertarian begs the question by assuming that the consensual transactions don't involve trading in things which someone else (say, the state) has a better claim to. This is not to say that the state actually does have a better claim—to work this out we will have to

consider difficult questions of moral philosophy and economics, *but merely to say that the bare fact that a transaction is voluntary between two consenting adults does not in itself establish its legitimacy*.

Indeed the "voluntary" arrangement might even infringe the rights of a party to the transaction. Suppose that you are the buyer, prepared to buy back your own television at considerable expense. In some sense you are a voluntary participant to that transaction, in another sense you aren't really a voluntary participant at all—you have a right to get the TV back either for free, or at only the cost of a finder's fee, presumably less than the present value of the television.

Appendix B: The tyrannical king as a benchmark

Suppose that a merchant went before a king and said "My lord, your taxes doth oppress me fiercely." To which the king replied "Begone vagabond! It is only by my clemency that I do not charge you far more! For this whole land is mine. Truly what you pay is not taxes, but a fee I impose on people who dwell on my property. What you think of as "your" land is really mine, and the land taxes you pay are but rent! It is nothing less than the natural laws of liberty which permit me to tax you."

From a deontic libertarian point of view, the only option here is to challenge the king's claims to own everything. I want to suggest that for many purposes we can use the king as a benchmark for assessing the plausibility of Libertarian defences of the status quo.

For example, if the merchant sayeth to the king—"But my lord—thou hast stolen all that thou possess, or inherited it from those who have!" and the king responds—"Certainly it is so—but further theft at this point to 'rectify' the situation would just perpetuate the cycle of stealing—best leave things as they are!" This argument would be transparently absurd, yet I've heard libertarians make exactly this argument for why we should not engage in mass redistribution to correct past wrongs, (and things libertarians regard as wrongs, such as government subsidies).

Appendix C: There are no golden strings, just institutions

Although this doesn't prove anything, I think it's useful to take a breath and clear our mind when we think about property. A lot of people imagine property as somehow metaphysically tied to a specific owner by intangible golden threads, and it's worthwhile to remind ourselves that this is not so.

Never forget that ultimately there are just objects. Tables, chairs, parts of land, and people, which are a special kind of object. What is property then? Property is a kind of social arrangement giving certain people certain bundles of permissions regarding certain objects, and denying those permissions to everyone else. In the final analysis then, like all permissions and refusals, *property is a collection of threats of social sanction, including violence.*

It seems deeply unlikely to me that we will ever be free of property understood in this way, or that this is even desirable. Even a communist state wouldn't want people trespassing in the nuclear power reactor without the right expertise—and what is the right to collectively exclude all people who lack special permission from a site but a kind of collective property?

Essential though it may be, re-framing property as the threat of sanction and violence, and not some metaphysical linkage, brings it into a new perspective. From this standpoint there is nothing especially 'non coercive' about, say, anarcho-capitalism, unless you take it as given that the claims it makes about who is entitled to what are ethically just.

Through-going subjective Bayesianism as a solution to the problem of scepticism

A common argument for the existence of God is that there is something rather than nothing. There are many good replies to this argument, but one of the more sophisticated challenges our sense that 'nothing' is the ordinary state of things and 'something' is an exception that needs explanation. There is an enormous, uncountably infinite plurality of ways the world could be. The 'nothing' world is simply one of these ways, and no more inherently 'natural' than the others. Why must the initial state of the universe be nothing unless we can give a reason otherwise?

We might talk then of the 'nothingness is natural' fallacy in relation to the initial conditions of universes. Scepticism seems to me to commit a variant of this fallacy in relation to epistemology. The set of permissible beliefs is assumed to start empty or near-empty, and then we have to fill it. A very low probability judgement for all statements is assumed to be the real correct default. Any exception is in need of justification.

There is an alternative view of epistemology that does away with scepticism. I see it as, broadly speaking, a development of G.E. Moore's argument against scepticism to a whole epistemology by way of subjective Bayesianism, though this might simply be

reading my own ideas into Moore's work. It is a development of subjective Bayesianism into a full epistemological standpoint, a kind of alternative to coherentism and foundationalism, or a dissolution of the questions they attempt to dissolve, depending on how you look at it. The view goes like this. There are no rationality requirements beyond consistency on the initial assignment of priors. Rationality only governs the *updating* or changing of our beliefs and does not impose any constraints on our *initial beliefs* beyond consistency requirements of the kind given by the Dutch book argument. The real province of normative epistemology is purely the assessment of changes in probability assignments. The demand for some ultimate justification of the initial judgements is an illusory over-generalisation from the practice of justifying changes in levels of belief. Our priors start out non-sceptical, and so our beliefs remain throughout a lifetime of updating—there is no higher ledge to stand on from which to critique this.

On the subjective understanding of probability, we do, after all, have to have *some* priors, and there is no known principled way of assigning them. Thus we arrive at some rather simple anti-sceptical proofs. For example, mirroring Moore's proof of an external world, we simply observe that our prior belief in an external world is rather high, and nothing has decreased it.

So why does the problem of scepticism appear to be a real problem? Over generalization. If I tell you that Susie is a lecturer in mathematics, you might ask how I came to know that, because you judge that it is very unlikely that it's simply built in to my priors that Susie is a lecturer in mathematics. This is true of basically all statements worth talking about outside a philosophical context, thus we get into the habit of thinking in terms of justifying our beliefs. When I then tell you that I believe there is an external world, and you ask how I know this, and I cannot give an adequate reply, it appears something is wrong. Really though this is a sort of illusion, all we are ever doing is justifying why we have decreased or increased our credence. Since your *a priori* belief that Susie is a mathematician is presumably low you need to explain why it has risen but the same is not true of belief in an external world, or causation- both of which start with high priors.

Carving up the philosophical terrain around personal identity a little differently

Many people are aware that there is a debate between the *psychological* and *bodily* continuity theories of personal identity over time. I want to carve up the logical landscape in a way which introduces a second, fully independent axis *substantive vs pattern* continuity. According to the substantive view of personal identity, you survive if whatever object makes you up survives. According to the pattern continuity view, you

survive if the pattern that you consist in continues, even if that means destruction of the object that currently instantiates the pattern which makes up you, followed by its replacement with an object that continues that pattern. The substantive view is often conflated with the bodily view, and the pattern view is often conflated with the psychological continuity view, but as we will see, such they are conceptually independent.

In order to demonstrate this, let's consider two classic cases which, when considered jointly, none of the four possible combinations (bodily substantive, psychological substantive, psychological pattern and bodily pattern) gives the same array of answers to:

1. **A brain transplant**

If your brain is placed in a new body and your old body is destroyed, have you effectively "changed bodies", or have you simply died?

According to both psychological views of identity (pattern & substantive) you survive a brain transplant in which your initial body is discarded. According to both bodily views of identity (pattern & substantive), you do not, since you are constituted by your whole body, not just your brain. So far we are in agreement with the standard account which treats all psychological theories as pattern theories and all bodily theories as substance theories.

2. **Teleportation**

If you are annihilated, then reconstructed elsewhere by a teletransporter, have you survived?

Here's where we diverge from the normal account. Teleportation is often thought to separate the bodily and psychological continuity theories of survival, with the psychological view contending that one survives teletransportation and the bodily view contending that one does not. Instead, in our taxonomy, what this case really separates is the pattern and substantive axis of views on personal identity. According to both pattern views of identity (bodily and psychological) you survive teleportation. In both cases the pattern or arrangement that, according to these views, consists in who you are, is continued, since the pattern of both body and mind is recreated. Also in both cases, the substance of what you are (either the actual mental states instantiated in the brain,

or the whole body) is destroyed, so both substantive views (bodily & psychological) rule this is a case of death.

Thus we see that the four possible combinations of views in our taxonomy are logically distinct, because none of them give the same answers as another to both cases above. The bodily substantive theory holds that one survives in none of these cases, the psychological substantive theory holds that one survives in the case of a brain transplant, but not teleportation, the bodily pattern theory holds that one survives in the case of teleportation but not a brain transplant and the psychological pattern theory holds that one survives in both cases.

Paradox of the book and the robot

This is an old puzzle—I recall reading it in Ted Chiang and I think in other places as well. It's a great example of how seemingly reasonable intuitions can lead us astray.

Premise 1. There could exist a book that contains infallibly accurate information about the future.

Premise 2. A robot could read this book.

Premise 3. The book might predict that, at some particular moment, the robot will perform some mundane action, like raising its grasper.

Premise 4. The robot might be programmed to be a perverse robot, in the sense that if anything or anyone makes a prediction about what it will do, it will do the opposite.

But it seems that premise 3 & 4 can't both be true, if the book is infallible and the robot has read it. It seems that, quite generally, if the robot reads the book, the book cannot contain any predictions about what the robot will do voluntarily (assuming the robot's programming remains intact and there are no errors).

But it feels weird doesn't it? Infallible future telling may not exist in our world, but it seems logically possible. If infallible future telling exists, there seems no reason why any agent shouldn't be able to access the results of that future telling, or any reason why it must suffer malfunction or abrogation of its programming if it does.

The logical paradox here is, at heart, related to, or even identical to, the grand-father killing paradox in time-travel. As there, we must say that in a world where time travel or

future telling is possible, any attempt to rewrite events (whether in a fixed past or known future) will always be thwarted.

New thought experiments for the backyard metaphysician to try at home

Sam and Finley

Finley loves Sam and Sam loves Finley. Sam hits their head, losing all of their autobiographical memory. However their skills and personality remain as they were. The usual question at this juncture is were you Finley, would you still love Sam? The idea being that if you answer yes, it follows that at least on an emotional level, you believe that Sam is still Sam. Thus, if your feelings are right, continuity of autobiographical memory is not required for continuity of personhood.

I want to ask a different question. It is harder to answer, and perhaps less philosophically illuminating, but still interesting. Suppose that you are Sam instead of Finley. You wake up and the concerned nurse explains to you many things—among them that you have a devoted partner that comes and visits every day. She gives some details of your life together.

My two questions are:

1. Do you think that, in this situation, you would immediately, or almost immediately, feel love for Finley—and not just the love you might feel for any kind stranger, but the love of a partner for a partner? If you wouldn't immediately feel love, how quickly do you think it might develop? How likely would it be to develop?

2. Regardless of your answer to the above, do you think you would be obliged to "try to love" Finley. Does the concept of trying to love someone even make sense?

The debate at the end of time

Everyone who has ever died has been raised from the dead in new and immortal bodies. Maybe your resurrectors used some of the technological options I discussed in "Oh Death, Where is the Antidote for thy Sting", or maybe they used supernatural power, it doesn't matter.

Your resurrectors explain that there is an important quandary—what should be done with the great wrongdoers of history? These have been raised alongside the rest of you. Does Idi Amin deserve an eternity in paradise, should Temujin break bread at the seats of the blessed? They have decided to leave these questions to a democratic decision of every human who has ever lived. A great debate begins, some arguing they should be absolved with everyone else, some arguing they should be imprisoned for a time, some arguing they should be killed(1), and some arguing for even worse. Who, if anyone, among the resurrected should be punished? How severely should they be punished if at all? How far does this go down the chain of wrongdoing? Should ordinary murderers be punished? Fraudsters? Those guilty of assault? Would you want cultural context to be accepted as a defense? As a mitigating factor? Would the sufferings the wrong doers endured in their own lives count as "time-served"? It falls upon you to take a stance on all these questions, or if not, justify your abstention.

Double trouble

Imagine a world where each body contains two persons each with a very separate personality. Which of these personalities is in the driving seat changes frequently. It's very common for one personality to be cruel while the other is kind, or honest where the other is deceitful.

How would you deal with punishment and criminal justice in this world, given that punishing a guilty person inevitably also punishes an innocent? What aspects would ethics require us to change? Assume you have a similar level of resourcing to a very well resourced penal system today. Would you try to make prison abolition work? Would you reluctantly accept prisons, but try to greatly minimize their use?

Conservation of moral status under misfortune

"His father was a drinker

And his mother cried in bed

Folding John Wayne's T-Shirt

When the swing-set hit his head"

–*John Wayne Gacy Jr*, Sufjan Stevens

I want to outline a (somewhat) rigorous way of understanding and defining "mitigating factors" when it comes to moral judgements of praise and blame, without appealing to any view in the debate on free will.

Regarding the serial paedophile, torturer and murderer John Wayne Gacy jr, the line in the above verse: "His father was a drinker" is an understatement. Gacy's father beat him, sometimes to unconsciousness. He suffered several severe head injuries as a child. These caused him to experience periodic blackouts throughout his adult life.

There is evidence of an association between both child abuse and head injuries on the one hand and serial killing on the other. This study: The incidence of child abuse in serial killers study on sexually motivated serial killers finds that while 68% of them survived child abuse, only 30% of controls did. Serial killers were also six times more likely to have been physically abused -like Gacy was—than controls. This study: Neurodevelopmental and psychosocial risk factors in serial killers and mass murderers reports on widespread head injury among serial killers. It is far from the first to do so. A link between head trauma and other forms of brain injury and serial killing seems to be folk wisdom in the field.

Two studies alone isn't definitive. But there is at least a prima-facie case for an association between serial murder, and misfortunes like these. Serial killers are a difficult population to study due to their rarity, so one option is to look at violent offenders in general. One imaging study High prevalence of brain pathology in violent prisoners: a qualitative CT and MRI scan study of violent offenders found they were five times more likely to have sustained a brain injury than non-prisoner controls.

One objection is that because the majority of abuse and head injury survivors don't perform horrific acts, we can't attribute Gacy's actions to these conditions. This argument confuses different kinds of enabling conditions. You couldn't take any random person, add these elements and have a killer—no one is claiming these are sufficient conditions. Rather, there is a very good chance they were part of what set Gacy on his specific path. Different people react to different stressors differently.

So Gacy was quite probably murderous due in part to the misfortunes he suffered. In me this evokes a sense of sympathy for Gacy, even as I remain appalled at what he did, and, to be honest, disgusted by him. Such arguments are well trodden ground. They are a favourite of everyone from philosophers to opinion column writers. They are often made in the context of arguments about free will. Bluntly, you've heard this all before right? Yada yada, tough childhood, yada yada, their behaviour was determined etc. etc.

Here's the (semi) novel point I want to make. I don't think you need to talk about free will at all to think these factors should affect how we see Gacy. In fact I think centring the debate on free-will muddies the waters, making it unnecessarily metaphysically sectarian. The bare fact that Gacy could have been someone different *and wasn't in part because of a misfortune that happened to him* moves me to pity—no metaphysics around free will needed. A lot of people I've talked to hold similar intuitions. This is important because debates over free-will are intractable. If we judge a certain class of people deserve clemency, then it's best if we can articulate this without appeals to our metaphysical views. Views on controversial questions in metaphysics like free will v determinism are hard to unite the public around, but, perhaps surprisingly, there is often much more common ground around ethics.

I suspect that the underlying ethical intuition is something like this:

Conservation of moral status under misfortune. *Consider the portion of the population who have blamelessly suffered some misfortune X. Call these people Xers and the people who have not suffered X, call them NXers. X hurts you and may cause you to behave badly. An omniscient, rational and morally good observer would regard NXers and Xers as overall morally equal on average, proportionally adjusting the praise and blame due to Xers accordingly.*

This seems to me to be a principle of justice. It holds because we should not, in general, blame the unfortunate more than the fortunate, or praise them less- to do so would be to morally favour people whose lives have gone well for them. We should accordingly adjust the "moral standing" of each Xer upwards. In some cases, like that of John Wayne Gacy Jr, this is still not going to be enough to get them out of the "bad" region, but it does make them less blameworthy.

We might even mathematicise it like follows. If a certain misfortune—like a head injury—makes you five times more likely to commit some horrendous act, that act is five times less blameworthy if you do commit it. So we have a (semi) formal basis for understanding mitigating factors. No misfortune should make you, in expectation, a more blameworthy person.

We have at no point referred to determinism vs free-will. The principle is workable—and attractive—even if you believe in the most demanding concepts of free will. This is so long as you accept the empirical premise people who suffer certain kinds of misfortune are more likely to do certain bad things. Of course the principle of the conservation of moral status under misfortune as I outlined it is far from complete and perfect—what if there were a kind of misfortune the suffering of which happened to correlate with

already being a bad person? Or how should we handle misfortunes for which people are partly—but only partly—to blame? However, it seems to me a reasonable starting point.

How to do things to words: mapping a post-analytic philosophy of concepts and intuitions

Introduction

I wrote my honours thesis on experimental philosophy, almost a decade ago. I then went on unsuccessfully to attempt a PhD. My feeling at the time was a feeling common to many philosophy undergraduates, but usually eventually beaten out of them. It seemed to me that many debates in philosophy were really, at heart, semantic or merely verbal debates.

A number of developments in the philosophical literature- from experimental philosophy, to the development of the idea of conceptual engineering (e.g. Chalmer's recent paper on conceptual engineering, which this post owes a great debt to and Haslanger's paper on race and gender) have led me back into this topic. I wanted to lay out, in simple English, a few thoughts I've been working on for years about the objectives of a post-analytic philosophy. Post analytic in the sense that analyzing concepts would not be a central objective, at least where analysis is conceived of in the normal way.

I lay out a number of ideas. Some of these ideas are mine, others are not. A lot of it it is stuff I've thought of independently, and later found others also thought of. The lack of citations does not mean I'm claiming credit, it just means that I'm lazy, and this isn't an academic paper.

A peculiar game

A big part of what we call analytic philosophy is the following game. I try to give necessary and sufficient conditions for a thing, and you find counterexamples which either A) exemplify the thing, but do not meet the conditions or B) Do not exemplify the thing but do meet the conditions.

Famously Plato proposed that man was a featherless biped and Diogenes responded by shaving a chicken

Let's suppose we are debating the definition of "life". I propose a definition- a set of necessary and sufficient conditions- life is any process that reproduces itself. You

respond by reminding me that fire reproduces itself, yet is certainly not alive. Perhaps I parry by adding to my definition that the putatively living thing must also achieve homeostasis. In this way the game continues.

Empirical evidence suggests that these games are interminable. No one ever wins. Even definitions that seem solid don't hold up. Knowledge as true, justified belief was one of the few "wins" of the methods, but as we'll see below, that was overturned by Gettier in the 20th century.

What we're going to do

In this essay, I want to lay out what I take to be three theses or phenomena that taken jointly undermine the tenability of this game as a pillar of philosophy, viz:

1. The stipulative theory of intuition

2. The family resemblance approach to concepts and

3. Diversity of intuitions.

Then I want to talk about some different activities which are an alternative to this game of trying to find necessary and sufficient conditions.

The stipulative theory of intuition

We'll use knowledge as an example. One (formerly) popular definition of knowledge is that it is justified true belief. That is to say, for something to count as knowledge you must:

A) Believe it B) Have an adequate reason to believe it (justification) C) And it must be true

This seemingly simple list of requirements held up as a definition of knowledge for almost two and a half thousand years.

However, there is a counter-example, owing to Edmund Gettier. Suppose Bob looks at his watch and it says 6:30. It is in fact 6:30. Bob's belief is true (because it's 6:30) and justified (because a watch provides adequate justification for a belief about the time). However unbeknownst to Bob, his watch is stopped- it stopped several days ago. It just so happens to be stopped on the right time by chance.

As a result of this counterexample, relatively fewish practicing philosophers today believe that to in order to count as knowledge, it is enough that a belief be true and justified. Additional conditions are thought to be necessary, although the nature of these conditions- and even whether there is a compact, simple list of conditions, is a matter of ongoing debate.

There are, I think, at least two important stories about what this Gettier intuition, and by extension, many similar intuitions about a huge range of different topics, are doing:

The theoretical account: On this account, our intuition that the Gettier case isn't a case of knowledge is like a hunch that a certain claim about the world is false. That hunch can be wrong.. Just like a belief that pyrite is a type of gold is wrong, so it could be wrong that Gettier cases aren't a type of knowledge. This hunch is thought to count as a type of evidence about the real nature of knowledge, for reasons that have never been entirely clear to me.

The stipulative account: On this account, our intuition that the Gettier case isn't knowledge isn't a theory about the world. Instead it's part of what fixes the definition of knowledge. Let's say that I define a Xuzazis as an object which is at least 50% green and weighs at least a kilo. I couldn't then find out that a Xuzazis can weigh less than a kilo. Similarly my "intuition" that the Gettier cases aren't knowledge is really me stipulating that any property whose extension includes the Gettier cases is thereby not the property of being knowledge.

It has long seemed to me that philosophers are insufficiently clear about which of these two accounts they think is true. I think this an important metaphilosophical dispute which deserves a lot more attention. Of course there are debates which relate very directly to this-debates over the causal theory of reference, over the Canberra plan etc., but it still seems to me that this is such a central question it should be discussed more, and more explicitly.

What I call the stipulative account has many advantages over the theoretical account. For one thing, it makes the epistemology of intuitions much less mysterious. The theoretical account has difficulty explaining how intuitions can be a guide to truth without resorting to mysticism. On the stipulative account there is no mystery. To the extent that intuitions are even beliefs at all on this account, they are beliefs that are their own truth makers.

For the sake of argument I accept the stipulative account as true, at least of very many important philosophical debates, for the rest of the essay. I acknowledge the possibility that the stipulative account is not true of everything. Philosophers have long noticed

that certain words seem attached fast to certain natural kinds- like the word “gold” and the chemical AU 79, but I think there is an important class of philosophical debates deploying linguistic intuitions which the stipulative account is adequate for.

Family resemblance

What if concepts don’t come with compact lists of necessary and sufficient conditions? What if what we’re looking for is more like a family resemblance? No one feature can rule you in or out.

One way this might work is through prototypes- the so called prototype theory of concepts. Maybe when someone says that X is a “bird” for example and asks whether we agree, what we do is compare X to a prototype bird- perhaps a hawk or a robin, checking for properties. The eagle passes fairly easily. The penguin leads to greater hesitation. The emu, even more so. There is considerable empirical evidence in favour of the prototype theory- although far from decisive. Note though that the semantics of words and the extension/intension of concepts can work like family resemblance even if the prototype theory of how we process concepts isn’t true.

Were inclusion in a concept more like family resemblance than a set of necessary & sufficient conditions, the classical philosophical approach to analysis would no longer be possible. Interestingly though, there would be an alternative, call it quasi-analysis. Quasi-analysis is the practice (purely hypothetical as far as I can tell) of laying out those features that tend to make something more X like- but not as necessary or sufficient conditions. So, for birds, a quasi analysis might include:

Wings, feathers, beak, lays eggs, can fly, builds nests, sings, squawks…

For knowledge it might include:

Is a belief, is true, is justified, is held very confidently, is justified by beliefs that are themselves knowledge, is widely agreed upon by those considered qualified to assess it, is certain, has been formed by a reliable mechanism, does not contradict other things the agent believes…

And so on…

Diversity between and within people

Experimental work on intuitions has revealed that there are systematic differences between intuitions about philosophical questions:

A) Between social groups (genders, races, classes etc.)

B) Between the same individual within different contexts (emotional state, disposition to blame etc.)

C) Idiopathically (two individuals in the same groups might have different intuitions just because- or maybe due to personality facets)

For example, coming back to knowledge again, experimental work on intuitions (sometimes called "experimental philosophy") has revealed that there may be both cultural and contextual influences on whether or not people consider the Gettier cases to be cases of knowledge. I believe also that there are variations along the lines of personality, and also of situation.

Result A makes it impossible to analyse "the" concept of anything, since there is not just one singular concept. Result B makes it unlikely that we can even hone in on a specific demographic group, and study their concept of X, because it is quite likely they have more than one, varying between situations. Result C puts a bow on top.

The joint effect of diversity between and within people, the family resemblance approach to concepts and the stipulative theory of intuitions

The joint effect of the three propositions I outlined is to make the game of hunting for necessary and sufficient conditions quite futile, although I don't think any of these points does it alone.

If just the stipulative approach to intuitions were true, and the other two propositions were false, we could keep hunting for necessary and sufficient conditions as an exercise in understanding the concepts in people's heads.

If just the family resemblance approach to concepts were true, it would be unlikely we'd find necessary and sufficient conditions that just so happened to meet all cases, but we might still learn some interesting things by playing the game, even if it were unwinnable. We might even say that, even if concepts work on family resemblances, it could turn out that some of them have relatively compact, workable necessary and sufficient conditions "by chance".

If just the diversity thesis were true, we might simply say that we had a lot more concepts to track- the different concepts of each culture, situation etc. We might even say -very brashly given the history of such things- that some cultures or people in some situations, had more correct approaches to certain concepts than others.

But I think the overall effect of these three propositions combined is to make playing the necessary and sufficient conditions game if not useless, at least of limited utility.

Once we accept these premises , what linguistic and quasi-linguistic philosophical tasks remain?

In what follows I lay out a number of linguistics and quasi-linguistics tasks that remain once we accept these postulates. The tasks I lay out are, at least broadly, philosophical tasks. My favorite task is at the end, so keep reading.

Concept creation (CC)

In concept creation as the name suggests we create a new concept. Probably ideally this is to go with a new word, but it might go with an old word as well, as a new meaning of that word. An example is my word Yvne, the inverse of envy. Yvne is cruel satisfaction that others are deprived of something you have, or have less of it than you do. Definition creation does not seem like an especially philosophical task on the surface, although on second thoughts finding blind-spots in our web of concepts and filling them maybe is a very philosophical thing to do.

Another great example of a philosopher creating a concept to direct our attention to something missed in ordinary thought- Tamar Gendler's concept of Alief's. To see what an Alief is, imagine standing on very thick and sturdy a glass floor over a deep ravine, going down hundreds of meters. You likely believe that you are safe, but you alief that you are not. Similarly if you are eating chocolate fudge shaped like faeces you likely believe that this is hygienic, but do not alief it. This is an example of philosophically interesting and provocative concept creation. [Chalmers seems to have thought of this as an example independently- too late to edit it out now]

Conceptual zoology(CZ)

There are a lot of already existing concepts of philosophical interest, waiting to be discovered by philosophers. Sometimes these exist as alternative uses of philosophically loaded terms- and thus have remained hidden from philosophers, who have seen them as deviant usages rather than appreciating them on their own terms. There is a lot of

work to be done discovering, classifying and understanding the role of such alternative concepts.

Consider, for example, what we might call the sociological concept of knowledge, commonplace among those who study the sociology and history of "knowledge" of various sorts. Here knowledge means something like socially sanctioned belief. Or at least this seems to me to be the definition at play. This concept of "knowledge" itself has various subtleties, and is worth the trouble to try to understand- and not just treat as a postmodern knockoff of the real thing.

We might also suspect that there is a scientific concept of knowledge. On the scientific concept of knowledge, a proposition can be "known" even if it is not really "believed" as such, or even true- it counts as knowledge just so long as we are justified in provisionally accepting it. We say that we know stuff to be true on basis of it following from relativity theory, even though it is quite likely that in a better, future science relativity will turn out to have been only a approximation. The proposition is thus unlikely to be true, not really believed, and only in a sense justified, yet it still would not be too strange to call it knowledge!

Conceptual redefinition(CR)

In conceptual redefinition I redefine a term for some purpose. The degree of redefinition can vary. I might try to capture what I regard as really meaningful about the term, or I might make something very different.

For example, "by knowledge, I mean justified true belief- inclusive of the Gettier cases" would be a conceptual redefinition of knowledge. A more radical reconstruction would be "by knowledge, I mean any correct belief, even without justification".

Here are some of the use cases for conceptual redefinition:

Social recognition: When gay marriage was still a goal, I would sometimes argue with conservatives who said that the common-sense definition of "marriage" was that it was between a man and a woman. Obviously I didn't accept this claim, but one of my favourite responses was that, were that true, we should change the definition for the sake of recognising an important group of people and their relationships.

Analysis: In the past I've suggested altering the term envy to include both what is currently called envy and what I call yvne. On such a redefinition, envy would be "a preference that others do poorly relative to yourself regardless of whether those others are currently above or below you". Such a concept, I think, would be useful for seeing

the world as it currently is. The current concept of envy is biased in that it focuses blame on those who are at the bottom of the social heap. In that regard it is ideological it represents the fear the powerful have towards their lessers, and conceals the truth that the rich can often desire the failure of the poor as much as the poor desire the failure of the rich. This is an example of championing a conceptual modification for purposes of clarifying analysis. In this example the analysis is social, but it could just as easily relate to the natural sciences.

Removing ambiguity: We can imagine a philosopher who, with a certain purpose in mind, declared that, henceforth by knowledge he would mean true, justified belief, even inclusive of the Gettier cases.

Normatively guided redefinition (NoGR)

This is a special case of conceptual redefinition where we try to make a definition correspond to a normatively significant category. Suppose I were trying to come up with a definition of "torture" for example, I might be focused primarily on a cluster of behaviours that are generally bad for the same reason. Maybe, for example, ordinary people don't use the word torture in such a way as to capture imprisonment, but I think imprisonment is in all morally relevant respects like paradigm cases of torture. Therefore I redefine torture to include imprisonment, on the grounds that this doesn't create distinctions without a moral difference.

The normativity doesn't have to be moral. Maybe I think that, although the Gettier case shows that justified true belief is not always knowledge. Nonetheless, I think justified true belief is always as epistemically praiseworthy as knowledge. I, therefore, propose that we should, either in a specific context or maybe even generally, redefine justified true belief as knowledge, because it matters and is valuable in the same ways that knowledge matters and is valuable.

Another potential example of normatively guided redefinition is the concept of survival as in that person survived thatevent. Or to put it another way, the temporal boundary conditions of the concept of personhood. For example, philosophers have long argued over whether one would count as surviving if one's body were disintegrated and reconstructed through a teletransporter. Increasingly an increasingly common view, argued by authors like Parfit (c.f. Miller for a similar position) is that this is the wrong question. Our intuitions about whether we survive this or that are hopelessly confused and unlikely to be turned into a single coherent narrative. Instead we should ask what do we care about? Mental continuity seems to me to be what I care about, regardless of whether you call this survival. Perhaps you are different though.

Natural kind hunting (NaKH)

This is a kind of extra-linguistic project that ties into the linguistic projects we're talking about here. According to the Stanford Encylopedia of Philosophy:

"To say that a kind is natural is to say that it corresponds to a grouping that reflects the structure of the natural world rather than the interests and actions of human beings[…]

Putative examples of kinds may be found in all scientific disciplines. Chemistry provides what are taken by many to be the paradigm examples of kinds, the chemical elements…"

In natural kind hunting, we look for natural kinds which share a similar extension to philosophically loaded words in our language. In some cases it might even be possible to find natural kinds which correspond exactly to our words. Historically this has often been done by people who think that natural kinds act like magnets for our words- but it doesn't have to be.

For example, I could uncover that there's a particular kind of brain state that corresponds to many, but not all, uses of our concept of belief. This would be a philosophically interesting discovery. We need not believe that it is revealing or changing anything about the definition of belief. Whether it does or it doesn't it is still, I think, an interesting scientific and philosophical task that relates to meaning.

To sum: NaKH might or might not be associated with a proposal to create a new concept which more precisely matches the natural kind, or with a proposal to reform an existing concept so that it matches the natural kind- but then again, it might not. Natural kind hunting is interesting in and of itself, and for many possible natural kinds (like those related to folk psychology- belief & desire), philosophers will have a lot to say in the hunt.

Philosophical lexicography (PAL)

We come to my favourite kind of project, which I call Philosophical Lexicography. Philosophical lexicography is a research program, continuing on from experimental philosophy, which aims to:

A) Map the usage and the variations in usage of philosophically important terms between groups of people and between the different contexts individuals find themselves in.

B) Understand these similarities and differences in terms of cognitive needs -universal and specific-, material circumstances -universal and specific-, personality factors, cultural factors, the history of ideas, the evolutionary history of our species, etc.

I have no doubt that this project of philosophical lexicography will be misrepresented as a relativist project- a kind of postmodernism in scientific garb. This isn't fair though. If Bob has a different concept of knowledge to Alice, for any given belief, B, there will be a fact of the matter about whether B is knowledge in Alice's sense, and a fact of the matter about whether B is knowledge in Bob's sense. There's no real relativism going on here. Different people mean different things by the same words, but we can hold the meaning fixed if we like, and there's only one reality that the words and meanings are being matched against.

Others will suggest that this project is all good and well, but that there remains a further fact about what knowledge really is, aside from our conceptions of it. I suppose there are ways this could turn out to be true but I see little reason to believe it, anymore than I find reason to believe there might be a xuzazis that weighs more than a kilo because there is a one true concept of xuzazis outside our heads.

Afterword for suspicious philosophers

I love conceptual analysis. I love playing with the intricacies of words. My own education and sympathies lie with the Canberra Plan. My real intention here is not so much to kill conceptual analysis, as to find a suitable afterlife. I've long disliked both the brash Quinean perspective of Epistemology Naturalised and the brash approach of trying to get intuition out of the picture by turning everything into a natural kind and combining it with externalist semantics. The kind of project I've outlined here leaves room for a paradise of Gedankenexperiment and counter-Gedankenexperiment, while not pretending that we're ever going to find necessary and sufficient conditions for anything unless of course, we declare them by fiat.

PART 5: MORALISM, IDENTITARIANISM AND OTHER MALADIES

Ugly, self-centred conversations

My least favourite kind of conversation is talking to people whose goals are centred on conversations. Some examples:

1. People who think good politics is centred on the right representations in discourse—use of the right shibboleths, avoidance of the wrong taboos- very common on both sides of politics.
2. People who think good politics is primarily civil politics, or the rare but real inverse—people who think one of the most important characteristics of good politics is that it be confrontational.
3. Leaving politics, people who view your conversational input primarily as a contribution to their clout or comment count.

We're not talking about people who think representations, civility, or clout have some value, we're talking about people who are obsessively focused on these. The problem is that when certain kinds of speech are themselves the goal, a cloying atmosphere is created. conversations become inward-looking. Everything becomes about *signalling* rather than *saying*.

There are many downstream effects, almost all bad. For example, conversations become emotionally disproportionate and unhinged, since very good or very bad things can potentially happen in the conversation itself—you're on the frontline and so you'll soon be shellshocked. Aggression is encouraged since the conversation is a battlefield for a better world or personal success.

The conversation is necessarily restricted because if it moves in certain directions that might itself cause the greatest possible evils, or thwart the greatest possible goods. All the energies of thought are being expended on policing language, rather than trying to say something interesting.

The ironic result then is that by raising the stakes of the conversation, this attitude makes trivial conversations. Language is a beautiful and meaningful thing, but it loses its beauty if it doesn't point beyond itself. The situation can become controlling, creepy and cultish.

Meanwhile, the real sufferings of the world continue unabated. I sometimes wonder whether people who focus on these things are even aware that thousands of people die or are ruined due to political choices every day.

It's like an ethical version of object impermanence. Everything that isn't immediately present in spectacular surfaces and discourse is discounted- in favour of gestures and tokens in a claustrophobic conversational game. Fuck the discourse, the spectacle or whatever you want to call it. Live in the real world, man.

Mistaken Identity and misunderstood interests: Haider and identity politics

I just finished "Mistaken identity" by Asad Haider, and like anyone who has just finished a good book I'm a proselytiser for it. My aim here is to draw out one thread of its multifaceted arguments, that the whole of the working class share a joint interest in abolishing racism in a way that is not recognised by what is often called identity politics. Like Haider we will only be discussing racial identity politics here, and focusing particularly on the problem of white supremacy in America.

Consider this quote from Ignatiev reproduced in *Mistaken Identity:*

"To suggest that the acceptance of white-skin privilege is in the interests of white workers is equivalent to suggesting that swallowing the worm with the hook in it is in the interests of the fish. To argue that repudiating these privileges is a "sacrifice" is to argue that the fish is making a sacrifice when it leaps from the water, flips its tail, shakes its head furiously in every direction and throws the barbed offering."

What Ignatiev is suggesting—and Haider concurs—is that rather than accepting that the white worker benefits in any unequivocal sense from racism we should consider a more complex view. As Haider puts it when discussing the history of racism and slavery in America:

“In exchange for white-skin privilege, the Euro-American workers accepted white identity and became active agents in the brutal oppression of African American laborers. But they also fundamentally degraded their own conditions of existence.”

In other words, Haider contends that white supremacy is ultimately bad for white workers and thus all workers have an interest in abolishing it, whether they do or do not recognise that interest. What I want you to consider is that this might be the real hinge of the whole argument over identity politics, insomuch as this is a useful debate. Do you accept that the best strategy for the whole proletariat, people of colour and whites alike, is to oppose racism, or do you believe that the objective interests of the white section of the proletariat lie in supporting and upholding white supremacy?

Identity politics comes from accepting the latter view. It takes both leftwing forms (moralism about race, hysterical demands that people must “renounce” their advantages etc.) and rightwing forms (fascism etc.) They are both two sides of the same coin-insomuch as they both accept that white workers win from racism, they simply differ in their moral assessment of whether this is good or bad.

1.

There’s a sort of cognitive trap here that it’s easy to fall into. It’s easy to think that obviously all whites have an objective interest in supporting white supremacy. After all, by definition, white supremacy is a situation in which whites are better off than non-whites.

To see why this doesn’t necessarily follow, consider a cute little meme that goes around the internet sometimes, usually an image or text that is a variation on the following:

“Bob gives Luke two cookies, Samantha one cookie, and keeps twenty seven cookies for himself. He then turns to Luke and says “Watch out! Samantha is going to take your extra cookie!”

(For clarity, this is not a quote from Haider)

Bob represents the largely white bourgeoisie; Luke represents the white portion of the proletariat; and Samantha represents proletarians of color. It is simultaneously true that Luke is better off than Samantha because of white supremacy AND that both Luke and Samantha would be better off were white supremacy dismantled. Luke is better off than Samantha under this unjust distribution of cookies, but would still benefit from, and be better off under, a more equitable distribution of cookies. To avoid confusion—this isn't just about money. Cookies also represent power, status, security and other goods. Whether in relation to material things or more intangible goods, a system that guarantees supremacy over another group need not guarantee a genuine improvement in living standards overall.

2.

What does the capitalist gain from racism?

Our first clue should be that the most open and vicious racists are also the most open and vicious in opposing workers rights. The struggles are linked if for no other reason than the clearest opposition to both is one and the same.

Racism is used by the right (yes, even the respectable right) to constitute an alternative *nexus of political struggle*. What I mean by this is that politics is always a struggle between forces, but even the nature of this struggle is itself subject to conflict. The right benefits from conceptualising the fundamental political struggle as a struggle between races and nations whereas the left benefits from conceptualising that struggle as a struggle between classes, in which racism is used as a crucial weapon by the ruling class.

Someone who sees politics as fundamentally a struggle between races will engage in cross class collaboration for the perceived advantage of their race. They may vote for those they see as having linked racial interests, but who do not share their economic interests, such as white supremacist capitalists, or rich persons of colour with bourgeoisie values. They may also be reluctant to collaborate on issues that should unite workers, such as joining a union that includes workers of color. Another example would be contemporary white workers who refuse to oppose police violence, despite police violence being a threat to workers everywhere, because they identify the police with their racial interests.

Thus racism creates a defanged and disorganised working class that doesn't even conceive of itself as sharing interests. It is a win for capitalists.

3.

Arguably at least, The mistake of thinking that white supremacy isn't ultimately counter to the aims of the whole proletariat leads into all the other problems with identity politics.

For example, a common complaint about identity politics is that it is moralistic and preachy, and that this leads to a culture of infighting and vicious online arguments. This follows from believing that there is no common linkage of interests in the proletariat to oppose racism. If you truly believe that some workers benefit in the long-run from racism, you will naturally resort to preaching and moralism, since the white portion of the working class has no material interest in abolishing white supremacy, there will be no option but to prick their consciences to guilt. Because of the range of many different ways people's interests are thought to be irrevocably divided- (gender, race, sexuality, ability etc.) it's no wonder that a low trust, high nastiness environment develops. Such an approach is very different from classical Marxism, which posits that almost everyone (except a small ruling class) is united in their interests.

Another complaint that's frequently made about identity politics is that it is overly liberal, and not sufficiently committed to abolishing capitalism. Again, this pretty plainly flows from the view that the working class is not even potentially unified with respect to its objective interests on white supremacy. It's a pretty major blow to a Marxist anti-capitalist view of things if the supposedly universally emancipatory working class has no basis for a solidarity of shared interests in opposing racism. If the working class is nothing special in this regard, a space is opened for class collaborationism of a liberal form.

4.

As we talk here about common interests, please keep in mind an important caveat from Haider:

"A common interest is constituted by the composition of these multitudes into a group. This is a process of political practice."

In other words, it is not so much that the working class already has a common interest in smashing both racism and capitalism, as that there is a potential liberating political strategy *which could infuse it with a common interest and a common program*. This is not a matter of a voluntary or subjective element deciding ex-nihilo to form such a coalition, rather it is a possibility already present in the class which comes to the forefront in certain circumstances.

5.

If the root of identity politics is a rejection of the revolutionary anti-racist potential of the whole working class as a group with a common interest in abolishing white supremacy, where does it come from?

This quote from Stuart Hall reprinted in Chapter 5, discussing economic despair in the 1980's and its interaction with racism in the UK captures it:

"As economic circumstances tighten, so the competitive struggle between workers is increased, and a competition structured in terms of race or color distinctions has a great deal of mileage. It is precisely on this nerve that the National Front is playing at the moment, with considerable effect. So the crisis of the working class is reproduced, once again, through the structural mechanisms of racism, as a crisis *within* and *between* the working classes"

In other words, as Bob gets a larger and larger portion of the cookies, the idea that both Samantha and Luke's situation could get better at the same time seems more and more distant. Since Stuart Hall wrote this, the wage share has been in almost continual decline in the developed world—for over three decades now.

Crudely speaking these material realities come to be reflected in our souls, or as Haider more eloquently puts it:

"I have come to think that this sadness is the primary cause of the restriction of politics to one's personal identity. Not only has the idea of universal emancipation come to seem old-fashioned and outmoded, the very possibility of achieving anything beyond the temporary protection of individual comfort seems like a delusion. Hence a call for universally beneficial social change is often heard as a personal affront: instead of an

affirmation of my individual demand for security and recognition, I am presented with a goal that lies beyond my powers to achieve."

We need to move our imaginations beyond the equitable distribution of crumbs, towards a coherent anti-racist, anti-capitalist program that roots itself not merely upon an abstract notion of 'social justice', but upon a recognition of our common interest in dismantling white-supremacist capitalism. Moving our imagination in this way is supremely difficult, because our despair is not merely a voluntary choice, but an outcome of our circumstances.

There are however reasons to hope. It is difficult to read the political weather, nonetheless it seems to me that over the last two or three years, an understanding of the inseparable linkage between anti-racist and anti-capitalist struggle has continued to deepen, one sign of which is the publication of the book that we have discussed here. As both capitalism and white supremacy continue to be exposed and treated with greater cynicism there is space for us to draw the links, and an urgency for us to do so.

On the perils of contrasting niceness with kindness

At the moment there is a low level cultural conversation going on about the difference between being *nice* and being *kind*. We are reminded that these things are different. It is possible to be a gruff old bastard with a heart of gold. It is possible to be sweet but selfish. I see this a lot on Twitter. Some people are even try to present them as opposites:

I am wary of this distinction for a couple of reasons. Let's start with definitions. By nice, I mean exhibiting a warm and friendly demeanor and being polite. By kind I mean being willing to make sacrifices on behalf of the well-being of others, and refraining from taking advantage of others. Here's why I don't think we should be so quick to pry niceness and kindness apart.

The statistical relationship

Open Psychometrics is a website that provides psychological tests people can take. You can opt in to make your results available to researchers, so I looked at the big 5 test they had—specifically the agreeableness facet ("agreeableness" is a composite factor including, among other things, both niceness and kindness). I took the two questions in the agreeableness facet which were obviously related, positively or negatively, to being nice, viz:

"I insult people."

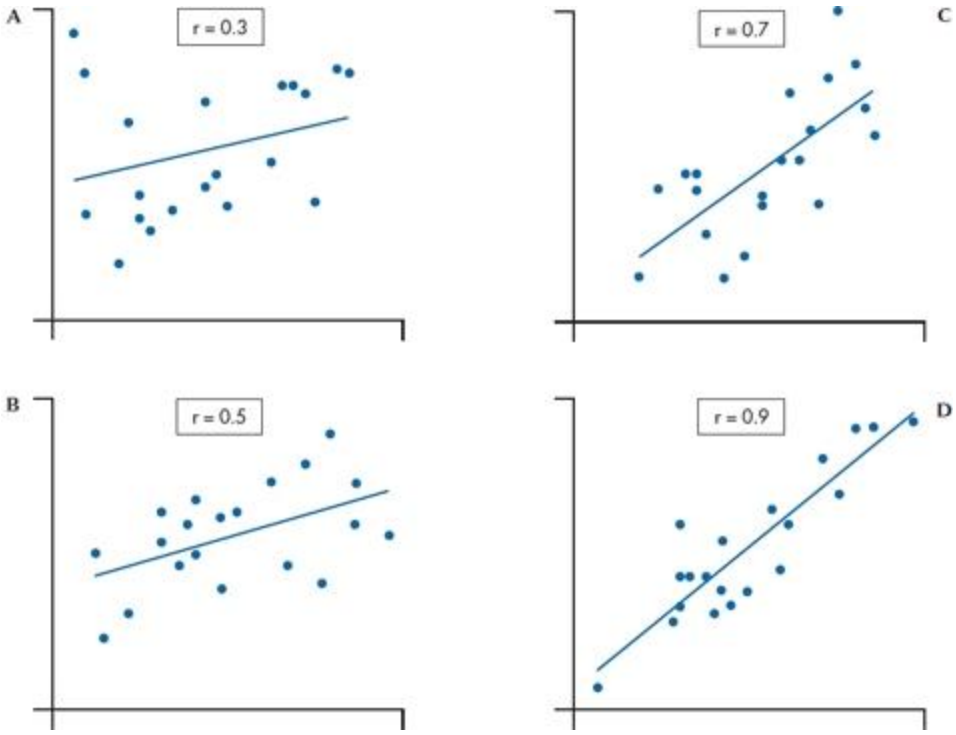
"I make people feel at ease."

And added up user scores for each (or rather added the question about making people feel at ease and subtracted the question about insulting people). Making a composite score of niceness.

I then took the three questions that seemed clearly related to being kind—positively and negatively. Subtracting "I feel little concern for others" and "I am not interested in others problems" and adding "I take time out for others". Making a composite score of kindness.

I then correlated the two composite scores, finding they correlated .37. This is a moderate correlation by the standards of social science, but the correlation is likely *attenuated* because the correlated scores are based on only a handful of questions—and very difficult questions to judge accurately about oneself regardless. When variables are measured unreliably, this makes the correlation between them weaker. This means the real correlation of the underlying variables, measured in a more reliable way, is doubtless much higher. You can read more about correction for attenuation here if you are interested: Correction for attenuation. I used some lowish estimates of the degree of unreliability and got a corrected score of 0.5.

For your reference, here is what different degrees of correlation look like graphed:



So while these variables may be, in theory, different things, in practice they tend to go together quite strongly. There is a robust tendency for the same people who are kind to be nice.

This is an important point that it's worth remembering in the social sciences (everywhere, really) just because two things are conceptually distinct as ideas, doesn't mean they easily come apart in the real world. The idea of having a kidney is very different to the idea of having a heart, yet every single organism with one has the other. For ease of mental labeling, you might think of this as the "conceptual distinctness fallacy"—the fallacy of thinking that the degree to which concepts are distinct from each other on a conceptual level determines how often the variables they track come apart in the world.

I would note one important caveat here. Be careful in applying this principle cross-culturally, because there are different norms of politeness across cultures. Just because someone comes from a brusque or abrupt culture doesn't mean they're ungenerous. The relationships between variables that exist at the level of individuals aren't always the same as the relationships that exist between variables at the level of groups.

So yes, there is a relationship between kindness and niceness, and that leads me to a further worry. I am concerned that people use this idea of "niceness isn't kindness" as a

crutch to excuse their own verbal cruelty to others. It can become a way to avoid difficult questions about how we treat other people. “Oh sure I told that person on Twitter to eat shit, so I’m not *nice*, but don’t you know I’m *kind* and that’s way more important”**(1)**. People can use this line of reasoning to avoid self-improvement. It could even be that if niceness and kindness are related on a deep level, letting our niceness atrophy may reduce our kindness as well.

But moreover, the effects of not being nice and being unkind are more similar than people realize—they can both fuck people up. Humans are fundamentally social creatures we are intrinsically—without our voluntary choice—effected by the opinions of others. Being insulted hurts. Being shunned or ostracized is *extremely hurtful* according to a wealth of psychological studies. One of the main, consistent results from the literature on ostracism is that being ostracized hurts a lot more than most people realize:

The Ordinal Effects of Ostracism: A Meta-Analysis of 120 Cyberball Studies

And these effects are particularly hard on vulnerable people—e.g. the socially anxious:

How long does it last? The persistence of the effects of ostracism in the socially anxious

So please, be careful in treating niceness and kindness as completely different, especially if your purpose in doing so is to diminish the importance of niceness.

(I always feel bad about moralizing so I want to be clear that I am neither especially nice nor kind)

---

(1)– Another notable feature of this is that our niceness is tested far more frequently than our kindness, so it’s easier to deceive ourselves about our kindness.

PART 6: FOR THE LEFT

Money and the Sceptic: A social-epistemological case for taking arguments for redistribution more seriously

The other day I saw someone put forward an argument against redistributive taxation based on a study, to which another commenter replied- 'but that study is from the Cato Institute'. Many people jumped in to object that this was an ad hominem argument, and the study should be evaluated on its own merits, Cato institute or no. This ordinary interaction on the internet got me thinking about the social epistemology of such claims and why I think we should not be so quick to dismiss ad hominem arguments in these situations, at least not entirely. Rather we must remain resolutely conscious of the social incentives shaping the discourse we are in.

Consider the following thought experiment:

You're aboard a spaceship and you crash land on a mysterious planet. To your surprise you discover an advanced technological civilization there.

Upon discovering the locals and achieving communications, you are asked to resolve an ancient dispute. You, they feel, are in an excellent position to be impartial. At first you try to decline the offer, but they are very insistent—the issue simply must be resolved one way or another.

They explain what the issue is. The juice of a certain fruit needs to be distributed between them. Ancient custom has acknowledged that for various reasons, some of them are entitled to vastly more juice than others. The issue to be decided is whether a more even distribution of juice should be adopted, or whether the customary distribution should continue.

The juice has the property of allowing its imbiber to go with less sleep than they would otherwise require. These creatures normally sleep for three quarters of the day, but those who possess the most juice can go for days without sleeping. This means that they can spend more time writing and discussing ideas. Very often the juice-rich write and think about the topic of the optimal distribution of the juice. They most often come to the conclusion that the existing distribution, or a slight tweak on it, is optimal—although a minority of them do support redistributing the juice. Sometimes they give out the juice to intellectuals of their acquaintance who then write about topics including the distribution of the juice. Since these intellectuals were selected by the juice owners, their writings most often conclude the current distribution of the juice should stay as it is.

Considering the situation, you find yourself with a problem. You suspect that the existing literature on the optimal distribution of the juice is probably quite biased, that results which seem to support its unequal distribution are found more often and receive undue prominence due to the structural advantages which support the status quo. However you are not an expert on the juice literature, it is very confusing and contains many mathematical symbols. There are a lot of stats, and every time someone puts forward a stat, someone else says that this is actually a misunderstanding.

What should you do?

We can debate exactly how much weight you should put on the lopsided origins of this society's thinking about the distribution of the juice. We would all agree though, that in the absence of the cognitive resources necessary to sort through the theories and arguments in ideal detail, if one has to make an assessment on the basis of one's limited information and intellectual powers, some weight should be placed on the reality that the side supporting the status quo commands greater resources, and this is bound to make its arguments appear stronger than they otherwise would.

Lifting the thin veil, if you haven't already guessed, the juice is money. The moral of the thought experiment is that in assessing the arguments for and against redistribution you should absolutely be acutely aware that there are not a lot of think-tanks funded by poor people. Even if you happened to be a trained economist, your mental capacities are still limited, and you do not have infinite time to check all things. You should, at a bare minimum, prioritise listening to and finding material by the side of the argument that is likely to be less well funded since you are less likely to encounter such material organically than you would be if both sides were funded equally.

Everything is negotiable on the right (and left)

I've heard people express the following astonishments:

"Why are conservatives arguing for everyone to go back to work? Aren't conservatives supposed to have stronger disgust and fear of disease instincts?"

"Why do conservatives so often viciously criticize Democrats who had military careers? Aren't they supposed to respect troops?"

"Why are so many conservative commentators criticising doctors and nurses online, don't they venerate first responders?"

“Why are conservatives tied to capitalism? Isn’t it corrosive of established traditions, family etc.?”

What they’re not understanding is that while individuals may care deeply about principles in politics, as a whole, politics is about coalitions not ideas. Ideas are wielded strategically on behalf of coalitions. The right represents a coalition of the powerful (and in our society, the core of that coalition is necessarily capital and capitalists). The left represents a coalition of the relatively powerless. Any principle claimed by either the right or the left will be dropped, at least by a majority, if it conflicts with the interests of their coalition.

This is why I’m on the left. On average, the demands of the powerless will be more righteous than the demands of the powerful. There are diminishing returns to status, money and power in terms of what it can do for you, thus those with relatively little status, money and power will benefit more from gaining some extra than the powerful will benefit from clinging to it, or grabbing more of it. If you want to get involved in politics with open eyes, the first step is to ask yourself not what ideas are more right but which groups, in the abstract, are more likely to have just claims.

If you want any leverage on historical processes you’ve generally got to cling to one coalition or the other and try to advance it and steer it. Sometimes you’ve got to try and steer it quite sharply, but not so sharply you’re flung off. As a rule, the world is governed by people, not by ideas.

A katana, an iron bar, and prison

1.

In the early hours of August the 10th, 2018, Jett McKee broke into the home of Hannah Quinn and Blake Davis. McKee pulled a gun (only later revealed to be fake) on Quinn & Davis and demanded money, he then knocked out Davis with a knuckleduster blow to the forehead.

McKee fled the home with Quinn in pursuit. Davis woke up, groggy, and chased after them both, carrying a katana. He then claims, not implausibly, that he found Quin cornered by the apparently gun-wielding McKee, and so he struck Mckee on the head

with the katana, killing him. There is some dispute over exactly what happened in the confrontation outside, with the prosecution saying McKee was simply running away. The judge's conclusion was that McKee had not posed any immediate threat to Quinn when he was struck with the killing blow. Whatever the jury thought had happened, they did not believe it qualified as self-defense. Thus, under section 418 of the NSW crimes act, the jury presumably either thought his conduct was not properly:

"to defend himself or herself or another person"

Or it was not:

"…a reasonable response in the circumstances as he or she perceives them."

The couple then panicked and went on the run before being eventually apprehended by police. Blake Davis was found guilty of manslaughter, but not of murder and sentenced to five years and three months imprisonment, with a non-parole period of two years, nine months. Hannah Quinn was found guilty of accessory after the fact and was given a community corrections order.

We'll give the jury the benefit of the doubt and assume manslaughter really happened and was proven here. The sentence of the judge was, in many ways, not unreasonable given prevailing legal standards (anything less might have been reversed on appeal), and very welcome given the almost comically aggressive way the prosecution pursued the charge.

Nonetheless, I think this is a great place to step back and think about why we send people to prison, and how we can reduce that number.

I won't bore you by reviewing in any great detail the case that can be made that Mr Davis shouldn't have been imprisoned but here is the case in brief: Whatever a reasonable response might have been to the circumstances, he had just woken up after being knocked out. He was presumably scared and confused, both because of the nature

of the situation, his own injuries, and his prior mental illness (PTSD). The character of his crime means he was very unlikely to repeat it. A severe non-custodial order would get the point across.

I want to change the way we look at prison. More generally, I'm concerned that we don't really *see* violence enacted by the state properly—it is veiled behind law and politics, and that makes us not properly understand it as what it is, violence just like any other violence. Like violence generally, it can be necessary, but before we can decide that it is necessary, we must do our due diligence by grasping it *as violence*. Once we've done that, it will become very clear that the use of prison needs to be restricted. What I'm going to propose is a sort of mental exercise, in the form of a thought experiment, intended to make us confront social reality without the veil of custom, depersonalization, abstraction, etc.

2.

Put yourself in the shoes of the judge. You've just decreed there is to be a custodial sentence. However, the officer of the court nervously clears his throat.

"Your Honour, I have just been informed by the prison administration that there is no room left in the prisons. However, our government, in their wisdom, have devised a cost-saving alternative."

The officer pauses and pulls out a large crowbar from a bag at his feet.

"Being beaten by an iron bar severely enough to be hospitalized is roughly as painful as a multi-year stint in prison. Therefore, her majesty's government has determined that in cases like this, beating with an iron bar will be substituted for imprisonment. Medical specialists will be on hand to carefully supervise the beating, to ensure that, while traumatic, it will not be lethal. Since you have assigned the sentence, you shall carry it out. Unless, of course, your honor would like to revise your sentence? Perhaps a large fine or a community corrections order with stringent restrictions would be sufficient?"

What I am suggesting is that if you wouldn't be willing to beat a person severely with an iron bar, you shouldn't be willing to send them to prison.

We're saying, for the sake of the hypothetical, that an iron bar is just as frightening as imprisonment (it's an equal deterrent), and incapacitates the victim as much as imprisonment from future crimes (say, by causing permanent damage to the musculoskeletal system). We're also saying that the permanent damage and difficulties caused are, on average, the same as being sent to prison.

Certainly, a look at the literature on the effects of imprisonment suggests that this doesn't make imprisonment out to be worse than it is. If anything, it grossly underestimates the effects and risks of prison relative to serious physical assault. Evidence from America suggests that in that country, one year of prison reduces life expectancy by two years, and that, depending on high-school diploma status and race, being incarcerated for the first time reduces lifetime earnings between 267,000 and 1 million dollars.

I'm not an expert on this literature, but I think both commonsense and survivor testimonies would tend to indicate that going to prison is extremely bad. If anything, it is quite a bit worse than getting thrashed with an iron bar.

I think that almost no one, save a few sadists, and perhaps some of the relatives and close friends of Mr. McKee would be willing to beat Mr. Davis with an iron bar. Now if Mr. Davis were, for example, serially violent, people might reluctantly agree to do it if the alternative was for him to simply walk free or face a comparatively mild community corrections order, but they would do so with "fear and trembling", conscious of the dreadful choice they were making, and of the impact, it would have on their own psyche.

My foundational claim—which is debatable, but which I will not get into debating here—is that the foreseeable consequences of your action are what you are accountable for. Perceptions of increased or reduced "brutality", abstracted from the actual

consequences of an action, are irrelevant to its morality. Above all, I'm saying that if directly viewing the consequences of your own actions, and having to cause those actions more directly would make you reconsider, *then you should reconsider without delay*. Don't send people to prison, or condone sending them to prison, if you wouldn't be willing to beat the shit out of them.

3.

What I want to rip through is something I call the *legal veil*. Our different reactions to prison vs beating someone with an iron bar are just one instance of the legal veil. Let me introduce it with an example.

Suppose a politician admitted, on tape, that the main reason they were introducing some bit of legislation that would jail thousands of drug users was politics. Oh, there were some considerations of public health, to be sure, but the overriding reason was to appease the police union and shore up hardline tough on crime voters.

There certainly could be a big reaction. It might even be talked about internationally. Very, very few people, however, would treat it as morally identical to the discovery that the politician in question was a serial kidnapper and torturer *even though they involve the same kinds of harms and the same kind of malice and aforethought*. This disparity in treatment is what I call the legal veil. We don't see officially sanctioned actions as equivalent to private action.

The legal veil has a number of different components to it. One is that legal actions are often, or perceived to be, "cleaner" and less brutally direct than things that do similar harm—hence the difference in our reaction to putting someone in prison and beating them with a steel bar.

But there's another component—the perception of *legitimacy* or *authorization*. I'm not denying legitimacy is real, I just don't see why it should excuse actions that are clearly not bona fide striving for the common good. This aspect, of the perception of legitimacy,

also relates to iron bars versus prison insomuch as if our society regularly authorized beating with iron bars, you'd probably feel a lot less squeamish about it.

There are some legitimate reasons why we should cut politicians, judges, senior public servants, etc. some slack. For example, people in these positions make an extraordinary number of decisions involving life and death and are therefore bound to get some wrong. My intention then isn't to excoriate for bona fide mistakes—the legal veil probably leads us to the right conclusion then in some cases, even if for the wrong reasons.

But often those who benefit from the legal veil are actually more culpable than those who do similar things for other reasons. Consider again, for example, the politician who extended the war on drugs for political reasons. He, unlike the man who kidnaps and tortures for his own sadistic pleasures, *has made specific oaths and promises not to do that sort of thing*. Arguably he's worse then. Now granted, if you had to pick one of them to babysit your kids you'd probably go with the politician, but this is a judgment of dangerousness, not of moral depravity.

Think concretely, specifically, and honestly about the exercise of power.

Should you care about that issue?

Does it:

1. Involve money or the economy
2. Involve the military or foreign policy
3. Involve criminal law. state violence or mob violence

If the answer to all of the above is no, it's probably a bullshit issue that's being used to distract you. Even if it does involve one of the above it may still be a relatively small issue that shouldn't be prioritised, but at least it's somewhat important.

This is not a plea for crude economism. For example, the right of trans people to access bathrooms is absolutely an issue, because it involves criminal law and violence. Immigration? Absolutely worth talking about under points 1 & 3. Racist policing? Absolutely worth addressing.

On the other hand, some idiot celebrity saying dumbshit on Twitter? Almost certainly not an issue. Arguments over Starbucks saying happy holidays rather than Merry Christmas? Not an issue. Trump looking like an idiot? Amusing, but not a real issue. Casting for a movie? Not an issue.

There are exceptions. Same-sex marriage was, I think, an important fight to win, even though the connection to any of those criteria is somewhat tenuous. An argument can be made that it involved 1, given the economic rights afforded to married couples, but even if these economic had been fully equivalent I still think the campaign would have been (somewhat) important. Nonetheless, you should start to get suspicious if an issue doesn't tick one or more of those criteria.

Thinking about political persuasion from a left-Wing point of view

1. The American left cannot win without persuading large swathes of the right & centre

There's a comforting lie that some parts of the American left like to tell themselves. We don't need to worry about convincing conservatives—we just need to get non-voters to vote.

This has never rung true to me. What evidence we have suggests turnout is not a panacea. For example, culturally the UK and Australia are very similar, however Australia has compulsory voting. The political centre of the UK and Australia is more or less the same despite this difference.

The evidence from the US suggests that non-voters in the US aren't as politically different from voters as is sometimes claimed. As of the time of writing, 53.5% of registered voters disapproved of Trump whereas when we look at polls of all adults… the figure is exactly the same—53.5%. Registered voters are more likely to approve of Trump than adults in general, but only very marginally (42.3% v 40.7%).

The Democrats would win if everyone turned out, but not by all that much. Specifically progressive and left-wing Democrats, even on the most generous conception of these,

would still be a long way from a majority. Thus there are strong reasons to think the left can't win simply by getting more people who share their values to turnout. Anti-parliamentarianism won't save you either—it's very hard to win a revolution if 75% of the population, at least, disagree with you. The idea that no persuasion beyond a little bit of base motivation is necessary is a comforting myth—a way of telling ourselves we don't have to talk with those self-satisfied, self-centred, self-serving, deliberately ignorant idiots over there.

There's a natural tendency to view those who disagree with us on topics which are genuinely important as abhorrent. In turn, abhorrent things are viewed as dirty, or likely to contaminate us. I'm not going to argue about whether these feelings are justified, instead let us just say they aren't useful—they're not workable levers for changing the world. If you pick a random person on the street it's almost certain that they'll hold extremely dangerous and regressive political views on at least one topic. I'm not talking about minor issues here—I'm talking about big things like war, criminal justice, etc. Despite that, it is absolutely essential that those who can engage with people and try to persuade them do so.

TLDR: examples from overseas, and data from the US itself, indicate that increasing turnout or motivating the base alone will not win the US for the left. There will be no left victory in the United States without persuading a lot of conservatives and centrists.

In the rest of this piece we'll go through the permutations and methods of persuasion from a left-wing point of view.

2. Arguing the line

The kind of persuasion that we are probably most familiar with is what I call arguing the line. Arguing the line is, quite simply, arguing vigorously for your position. Sometimes this is done against a real interlocutor, as in a comment or Twitter thread, and sometimes this is done against a purely hypothetical interlocutor, as in many blog posts. Arguing the line is not a collaborative process, it is a confrontational process, although it is not *necessarily* cruel or angry.

Some would say that this is the least effective of persuasive strategies; I disagree, although it is often overplayed. In order to see how arguing the line can be effective, it's important to understand what it will generally **not** achieve.

Usually, arguing the line is not going to change the position of the person you are arguing with on the spot, especially if the argument is happening on the internet. If it does change anyone's position immediately, it will almost invariably be on small points.

Rather than changing the mind of the person you are arguing with, the primary purpose of arguing the line is to convince onlookers. There are a lot of people with relatively unformed political views floating around in pretty much every space on the internet. If you're on the fence, seeing someone argue coherently, reasonably and powerfully for a position like Medicare for all or an end to foreign interventions can have a big impact.

Keeping in mind your real audience—undecided observers rather than your direct interlocutor— clarifies the mind. It will help you pick your battles, keep your morale up, and refine your methods and pitch. This isn't to say you should just speak as if your direct interlocutor weren't there or isn't worth paying attention to—in most contexts this would make you seem weird, rude or aloof.

3. Rules of thumb for arguing the line

You should aspire to state your arguments so clearly that no one can misinterpret you even if they want to. This is because it is quite likely your opponent will be deliberately or quasi-deliberately trying to misinterpret you. You almost certainly won't succeed in making your work impossible to misinterpret, but it's important to try and get as close as possible. This is because if you're engaged in a back and forth with someone, onlookers will only be partially paying attention. Thus if your opponent attributes a meaning to you, many onlookers will automatically assume their interpretation is correct unless you have been so totally clear that even people who are only half paying attention can see that your opponent is bullshitting.

Often people's impression of the epistemic virtues of the debaters stays with them longer than their recollection of the actual arguments (e.g., "This person seemed reasonable" vs "This side seemed histrionic or dishonest.) Thus, without seeming like a pretentious dickhead, make your epistemic virtues visible. Show others that you are measured, calm, inquisitive, nuanced where nuance is appropriate, perspicuous and attentive to the whole picture. If you aren't already these things—try to be! If you can write or speak well, do so.

A good rule is that you should avoid engaging where you are clearly going to get stomped. This includes topics where you have no idea what you are talking about and circumstances where your opponent can control the flow of the conversation in such a

way that they can cut you off at leisure. There's an old proverb about this, it's harsh but it makes its point: It is better to remain silent and be suspected of being a fool than to open your mouth and remove all doubt. The point being that if you don't say something your side will be perceived as having lost ground, but not as much ground as if your opponent can smash through a tissue thin defence.

Consider the way Ben Shapiro bolsters the rhetorical strength of his case by picking dissenting audience members currently under the grip of strong emotions, controlling the flow of conversation and "destroying" them. This is a great example of why it is sometimes better not to engage if you can't do so on fair (or better than fair) terms.

The above rule has to be tempered with the recognition that there is sometimes value in being the lone dissenter. If you are the lone dissenter, you're certainly going to 'lose' the debate, since the numbers of the other side mean they will get more speaking time—they can 'rebut' all your points and put forward more ideas than you can reply to. Nonetheless, there can be value in clearly, simply and powerfully stating your ideas. In the Asch conformity experiments—for example—suggested that a group consensus about something has an extremely powerful effect on onlookers, but even a single dissenter can greatly weaken the effect of that conformity. We can think of this as the principle of contested space—if there is a space, conceptual or physical, which the left is not contesting to at least some degree then there is a problem.

(Incidentally, fuck the left-wing purists who will have a go at you for entering, participating in and contesting non-leftwing spaces, they're among the very worst the left has to offer. Which is not, of course, to say that you should be posting on fascist boards)

People—even fellow travellers—always try to pigeonhole arguments into being a variant of something they've already read—either to dismiss it or accept it without thinking too deeply. People are always looking to be able to say "oh this writer is one of X type people so she thinks Y&Z and must be vulnerable to objection P". In order to avoid this, try throwing in curve-balls that will surprise your readers expectations of what they think you believe. For example, taking a somewhat trite example and channelling the Communist Manifesto for a moment:

"I'm a Marxist so I believe that capitalism has accomplished wonders far surpassing Egyptian pyramids, Roman aqueducts, and Gothic cathedrals; it has conducted expeditions that put in the shade all former Exoduses of nations and crusades."

4. Rogerian persuasion

If you want to actually persuade an individual of something, and not just onlookers, Rogerian persuasion (named after the psychologist Carl Rogers) is your best bet.

Most people don't have especially clear or fixed views on issues, but instead have a mixture of beliefs and values related to any given topic. The idea of Rogerian persuasion is that if you want to persuade someone on any given topic, you should focus on areas of shared and similar beliefs and values. You want to demonstrate how those beliefs and values might actually support a left-wing position on the topic. This is easier and less artificial than it sounds, because most people at base have many quite left-wing intuitions and beliefs, they just get crusted over by reactionary propoganda.

Focus on demonstrating that you understand what the other person is thinking and saying. A good technique to combine Rogerian persuasion is what counselors refer to as mirroring. Paraphrase key things the other person has said and repeat it back to them to show you understand and check that you are on the same page.

One very important point in Rogerian persuasion is never to leave the other person in a position where they don't have an out. You want them to have a natural route of escape. A way they can walk back from positions and change their mind without making a big mea culpa. People usually aren't that afraid of changing their minds, what they care about is the humiliation of having to admit that they were previously wrong, especially if it is in a way they now recognise is a bit repugnant. As a result, people often want to dress up a big change of heart as simply stating something they've always believed 'more clearly' or 'clarifying' their views. There's a fine line here. I'm not sanctioning dishonesty, and there probably are times when people should feel a little bit uncomfortable. But remember, this isn't about 'winning', much less about punishing the person for their prior views. It's about the transformation of the world.

Don't try to turn Rogerian persuasion into passive-aggressive hippie focus-group bullshit where you get exactly the cookie-cutter result you want. You really do have to listen, you really do have to actually care what the other person thinks and accept—at least in the context of that conversation—the differences in your opinion. Above all you have to respect the autonomy of the other person. This respect for what the other person thinks means that you're not going to turn out intellectual clones of yourself, but that's okay.

Sometimes you've got to accept partial wins. For example, if you can persuade someone who supports the death penalty, to restrict that support to a much smaller set of circumstances, that's a win. If you can persuade someone to move from supporting the criminalisation of abortion to just being personally opposed, that's a win. Accepting these partial wins does not mean having to compromise your own views.

5. Mere presence

In a lot of ways, this is related to Rogerian persuasion, but it's worth emphasising separately. Simply being a part of someone's life while holding left-wing views can exercise a powerful influence.

Just letting others know that, for example, you support free public college tuition has an effect. You are giving the other person information—that it's possible to be a reasonable, kind person and believe these ideas. For a lot of people exposed to an intense diet of right-wing memetics this is a powerful thing, since their understanding of the world includes the assumption that it's only weirdos who think those things. Try letting people know you're left-wing, being a presence in their lives, but also being cool about it.

6. Don't forget the Socratic Method

Socratic questioning is a kind of arguing by question, where rather than concentrating on putting forward propositions of your own, you focus on asking difficult questions about what the other person believes. In the ideal case (as Socrates practiced it) Socratic questioning leads the other person to move to your own position, as they struggle to deal with the difficulties you raise by amending their position step by step. Even if you don't get that far, Socratic questioning is a powerful method. Socratic questioning can complement either Rogerian persuasion or arguing the line, although the kind of Socratic questioning that works best will vary depending on your purpose.

Intuitively it can look like the person asking the questions is less powerful than the one giving the answers. It's the person answering the questions who gets to describe their worldview, and who speaks the most. This is an illusion however; there is immense dialectical power in asking the questions. When someone is simply expounding their view they can make big logical leaps which are all to easily concealed from the casual reader. Under questioning though, this stuff comes out. You can really expose the underlying assumptions.

Here's a great example of what Socratic questioning can look like, owing to Current Affairs magazine podcast:

"Single payer can work in places like Sweden because they are more homogenous, the United States is too diverse for single-payer healthcare.

"Okay, Canada is only a bit less diverse than the US. What do you see as the key differences between diversity in the United States, and diversity in Canada, which makes single-payer possible in Canada but impossible in the United States?"

The question sounds very innocuous, but is actually quite difficult to answer without either A) implausibly insisting that the relatively small quantitative difference in diversity levels makes a huge difference. B) Saying that the problem is the kind of ethnic groups the US has—straying dangerously close to explicit racism C) Just outright changing the topic.

7. Make propoganda

I don't have much to say about this except an exhortation: Make and distribute stuff that can persuade people: memes, posters, pamphlets, wearables, comics, drawings and essays.

The majority of internet users (as around 99%) are largely passive. Outside the internet, the ratio of culture consumers to culture producers is even higher. You really don't have to try very hard to have an out-sized impact (hundreds of times that of the average person) on the conversation. Look at what other people are doing who are good at making persuasive political materials, study their technique, experiment and, hey presto, you'll almost certainly find there's at least one medium where you can excel.

8. Organising as persuasion

It's a pretty well known observation that the process of fighting for justice is radicalising. Thus if you want to persuade people to the left, you should start organising. The reasons being part of organising tends to draw people to the left are many, but include:

A) The support they will (hopefully) receive from the leftists.

B) The conversations they will have with other people they are organising with, and the shared concerns and experiences they find together.

C) The opposition they will face from capital and the capitalist state.

There are limits here. For example, around the world numerous farmers have been organised to oppose fracking on their land. While this experience has no doubt moved the campaigning farmers to the left in some ways, in many places the majority of these farmers will still vote for centre-right parties. The limits are, based both on the objectives of the campaign, and the class and social position of those participating. Nonetheless, organising changes people.

9. Institutions as persuasion

Left-wing institutions are the useful residual of concrete left wing struggles and organising. For example, many unions can trace their existence, however distantly, to a particular wildcat strike. Unions are the ultimate example, but not the only one, even within the sphere of industrial issues. For example, although they are rare in this period, it was common in the past to have worker's education institutions, workers schools etc. Most of these can trace their origin to some particular flare-up in the worker's struggle. The same is true of women's libraries, associations of racial minorities, pride marches, even the much maligned student union. These institutions often owe their existence to big moments in particular fights, and while the struggle continues, they often outlive the specific campaigns or moments of intense action that gave birth to them.

I'm including them in this guide because these organisations perform persuasion on an industrial scale, they aim to align not just individuals, but whole demographics and suburbs to a cause. Their strategy is a form of persuasion, but it transcends persuasion, when successful they create whole new political categories and identities.

One of the major problems with sectarian organisations is their tendency to take for granted these kinds of institutions and not recognise their value except as a momentary tool of the sect.

Inversely though, it would be a mistake to regard these organisations as inevitably radical—they tend to become liberal over time when disconnected from struggle. Too much faith in these organisations is linked to that common new-left disease, the tendency to venerate oppressed communities without recognising the contradictions that exist within such communities.

10. A word on critical thinking and informal fallacies

Many Universities have courses on critical thinking. In a good critical thinking course one learns about formal and informal fallacies, cognitive biases, the scientific method, the basics of probabilistic reasoning sometimes up to Bayes' theorem, a tiny bit of formal logic, maybe a smidgen of inferential statistics, and a few other useful tidbits.

A lot of this material, but especially the study of informal fallacies has been given a bad name by poorly socialised people who try to use it like incantations from Harry Potter ("Ad Hominem!","Petitio principii!") and don't pay attention to the larger conversational context. Fragments of reasoning that would be fallacies in one context are perfectly valid in other contexts. Sometimes arguments that appear to contain informal or formal fallacies are just abbreviated statements of perfectly fine arguments.

My advice would be to familiarise yourself with ideas like informal fallacies, cognitive biases, probabilistic reasoning etc. but generally don't use the words and terms in your explanations of your thinking. Instead explain the basic flaw in your opponent's reasoning without appealing to the jargon of cognitive biases or informal fallacies. There's two good reasons for this. The first good reason is that you should be avoiding jargon generally. The second is that you'll avoid the bad reputation that these particular conversational manoeuvres suffer.

Specifically with regards to ad hominem attacks directed against yourself, either ignore them, or, if you must, retaliate with a similar insult or comeback. Pretty much everyone understands that ad hominem quips don't really prove anyone right or wrong. Complaining that your opponent's insults are fallacious won't do you any favours and just comes across as whining.

11. Dealing with bad faith

A lot of people don't want to engage in persuasion because inevitably many of those who wish to discuss politics are acting in bad faith. This is a serious problem, the only advice I have is try to make careful and reflective judgement calls on when discussion is worth your time.

For example, there's little point arguing with someone who clearly isn't willing to listen if there isn't an audience of potentially undecided people to see your argument (although, since the majority of posting is done by a small minority of people, the probability of you having an audience is usually higher than you think.)

In general, beware time wasters, but recognise that on occasion time wasters will successfully waste your time, and this probably can't be helped.

The far-right are a special case. Arguing with the far-right has many dangers and few benefits. For example, some ideas are so niche that they gain relatively more oxygen if you argue with them even if you completely squash it. Let's say someone comes up with some novel far-right position or titbit and you completely squash it. Good for you, except no one had even heard of it before you bothered so no matter how thoroughly you squashed it, you've now helped it enter the discourse.

There's a special kind of bad faith associated with far-right argumentation. As Sartre puts it in relation to anti-Semites:

"Never believe that anti-Semites are completely unaware of the absurdity of their replies. They know that their remarks are frivolous, open to challenge. But they are amusing themselves, for it is their adversary who is obliged to use words responsibly, since he believes in words. The anti-Semites have the right to play.

They even like to play with discourse for, by giving ridiculous reasons, they discredit the seriousness of their interlocutors. They delight in acting in bad faith, since they seek not to persuade by sound argument but to intimidate and disconcert. If you press them too closely, they will abruptly fall silent, loftily indicating by some phrase that the time for argument is past.'

In other words, if your opponent has bought into the asetheticisation and/or gameification of politics, and cares not whether they are right or wrong—but only for power—why bother? Prove their thesis that their ideas will give them more power wrong in practice, by usefully spending your time elsewhere.

The only thing I would caution here is that you shouldn't use not talking to the far right as an excuse. There's a sense for example in which what the typical Trump supporter believes is far-right by many reasonable standards. However, applying such a definition in an American context simply wouldn't be useful. If you're not comfortable talking to Trump supporters personally, fine, but don't make a principle out of it.

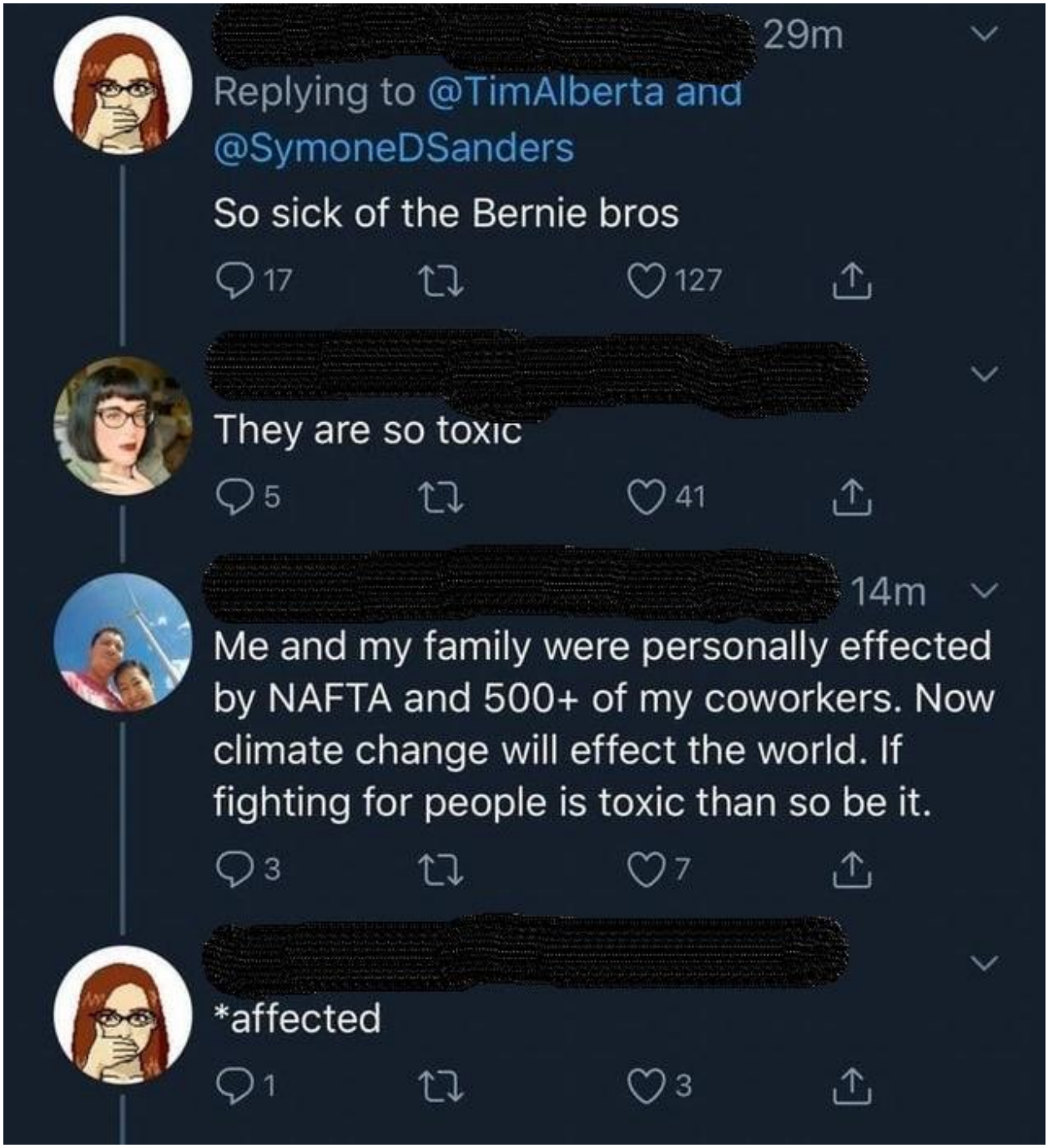
I don't know how to tell you that politics is about murder

There have been numerous blow-ups during the Democratic primary about words and behaviour. The interminable debates about Joe Rogan. Elizabeth Warren partially attributing her non-endorsement of Sanders to snake emojis. Endless tone policing of Bernie surrogates like Briahna Joy Gray. The details have been stripped like desert bones and are, in any case, irrelevant.

I'm concerned about the enthusiasm with which these arguments have been taken up, and I think you should be concerned about this too because the substance of politics is murder. It's about the avoidable deaths—and the enslavement—of human beings. Any discussion related to politics that is not directly related to life and death, or human immiseration, should be of secondary importance. Almost 1% of the US population is incarcerated. 26,000 Americans die a year due to lack of health insurance. There aren't even good figures on the number killed by American sanctions each year.

Maybe this all sounds a bit repetitive, but sometimes you've got two options, silence and repeating yourself about a message that already should be been heard, and it's better to repeat yourself.

This isn't an essay to tell anyone they must believe as I do. Maybe voting for Joe Biden is the right way to stop the ghastly procession. I do not accept as legitimate however the idea that we should decide that on the basis of this bullshit. If you are making your choices on this basis or pretending you are, you are doing the wrong thing and must stop. The questions of power cannot be evaded. Total political apathy would be more honest and clean.



It happens on the left too. For example, Mike Bloomberg's personal behaviour has often been appalling. Nonetheless, it bothers me a lot that we paid more attention to his personal cruelties than to the thousands of people who died avoidably, and the hundreds of thousands who were harassed by the police unnecessarily, during his mayoralty. Why are we drawn back into these personalities all the time, it's not even like they're particularly interesting personalities.

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

Actions that unnecessarily kill people don't somehow take on a different moral meaning just because they're done by a politician *Thumps table* *Thumps table*. The inability to grasp that murder through politics is no different to hiring a hitman really colours our perception of the moral realities of politics.

People have a great deal of trouble giving murder and wrongful death the attention it deserves unless it happens in a spectacular, made for television way. Our natural priorities when it comes to politics seem to be:

1. Big events with relatively small death counts (e.g. 9/11)
2. Meaningless celebrity gossip type bullshit
3. Stuff that actually kills a lot of people (sanctions, lack of healthcare etc.)

This tendency has long worked against justice. Mark Twain remarked of the French revolution:

"THERE were two "Reigns of Terror", if we would but remember it and consider it; the one wrought murder in hot passion, the other in heartless cold blood; the one lasted mere months, the other had lasted a thousand years; the one inflicted death upon ten thousand persons, the other upon a hundred millions; but our shudders are all for the "horrors" of the minor Terror, the momentary Terror, so to speak; whereas, what is the horror of swift death by the axe, compared with lifelong death from hunger, cold, insult, cruelty, and heart-break? What is swift death by lightning compared with death by slow fire at the stake? A city cemetery could contain the coffins filled by that brief Terror which we have all been so diligently taught to shiver at and mourn over; but all France could hardly contain the coffins filled by that older and real Terror — that unspeakably bitter and awful Terror which none of us has been taught to see in its vastness or pity as it deserves."

My crackpot theory is that it comes from a kind of illusion where some people believe that there are only two kinds of people in the world—people they know, and celebrities (inclusive of politicians). This is because these are the only people we regularly encounter.

A less kind theory is that this gossipy rubbish is conscious or semiconscious dissimulation. People talk about this stuff because if they talked about the substance of politics when defending their choices, they'd have to admit they aren't half as left-wing as they pretend to be.

A brief note on the disposability ideology

Consider three areas in which America is uniquely bad:

Criminal justice

*America imprisons more than twice as many people per capita as its nearest developed competitor, Israel.

*The US imprisons children at a rate of 60 per 100,000, 11x higher than the western Europe and the highest in the world.

Healthcare

*The US is the only developed country in the world where people often die because they don't have enough money for medical treatment.

Workplace relations

*No other developed country has at-will employment.

There is an underlying ideology that unites these- an ideology of disposability. If people do the wrong thing or don't produce enough value, it's alright to get rid of them and it's not required to try other things first. While the disposability ideology it isn't unique to America, it takes its most perfect form here. The areas we've listed mark the biggest departures of US domestic policy from other nations. Thus, we could call US capitalism "disposability capitalism". All capitalism treats people as disposable, but the United States has perfected it. Economic security, liberty and life itself can be discarded when they are inconvenient.

Even Americans who affirm that people are not disposable will make exceptions in practice. For example, murderers, rapists, violent people, and people they disagree with politically. The left may be less vulnerable to disposability thinking, but they are not immune. None of this is to deny legitimate communal-defence, but everyone seems quick to draw that blade.

The reply to the disposability ideology is often to argue that this or that person or group didn't deserve to get junked. It is true that people's sins are often less than claimed, but

it is not enough to say this. It might sound trite, but we should affirm that everyone deserves dignity through the bare fact of their existence.

For communism and against foreclosure on the future

By communism here we mean a system in which the principle of:

From each according to their ability, to each according to their need

Is implemented as the sole principle of economic distribution.

There has never been an advanced communist society—only societies that aspire to communism in the long run. Money still existed in the Soviet Union, China and Cuba. The state paid some people more than others, and not on the basis of their extra need. It is not clear that there has ever existed an advanced society that even had the capacity to build communism.

There are I think two main objections to the possibility of a technologically advanced communist society, viz:

1. The incentive problem
2. The calculation problem

The incentive problem is the problem of making sure the work gets done, especially boring, dangerous and stressful work. If everyone gets paid on the basis of need rather than effort, why would anyone want to do these kinds of work?

The calculation problem is more technical, consult the Wikipedia article “The Economic Calculation Problem”.

These are legitimately tricky problems. One way to respond to these difficulties is to give up on communism—for years I did just that, in favour of social democracy or democratic socialism. Another response comes, ironically, from a right-libertarian I knew in University.

II.

I was putting to this libertarian the technical problems with anarcho-capitalism. These include the provision of public goods and management of externalities. I argued that these problems showed anarcho-capitalism was either impossible or undesirable. His response to this was twofold:

1. A) Anarcho-capitalism is a regulatory ideal—an organising concept for political action. Something can serve as such an ideal even if we don't yet know if it is possible.
2. B) The world we live in would have been unimaginable through most of history. To foreclose on the possibilities of the future would be foolish. Tremendous changes in technology and social institutions await.

Although I'm certainly no anarcho-capitalist, I think this is not a bad defence of aspiring to anarcho-capitalism, despite the technical difficulties.

These days, this is how I feel about communism. There are forces that push in the direction of communism. Communism is a good regulatory ideal for those forces. It unifies the most radical progressive tendencies in society. Will it ever happen? Predicting the direction of social and technical advances in advance is impossible so it would be foolish to rule it out. Insomuch as it centres human needs in the productive and distributive process it is a good *aspiration*.

Note: Consider especially advances in AI and transhuman enhancement.

We need to be exact about what is being proposed here. I'm not arguing "No one can prove communism is impossible, hence we should believe that it is possible." That would be to substitute faith for reason. Instead, the idea is that "We don't know whether communism will ever be possible but it may well be. Moreover there are reasons to think it is desirable, and we can take steps to try and approximate it better in the present, in this sense of aspiration it is right to aspire to it.

The danger of an aspiration like communism is that it can overwhelm more specific social plans for the coming decades. So long we avoid this danger, we should happily regard ourselves as communists.

III.

I wanted to talk about another issue which is quite logically distinct, but which I regard as emotionally linked: reform vs revolution. What I say here will not be new to anyone who has thought about this for a few years, but I remember that there was a time that I didn't understand it, so maybe it will be helpful for some.

There is no absolute divide between revolutionary and reformist strategies. There are only degrees.

The most obvious reason this is true is that winning reforms can increase revolutionary morale and organisation. Conversely, the threat of revolution can win reforms.

The deeper, often missed point is that winning elections can create legitimacy for an anti-capitalist movement. If the deep state then responds through a coup, a revolution can happen in the defence of an elected government—revealing and hopefully defeating the true nature of the repressive apparatus(1). If successful, such a revolution can then dismantle the existing capitalist state. For this reason, revolution and electoral victory shouldn't be counter-posed strategically, at least not in an absolute sense.

Engels describes how revolution can be a defensive option, once the state represses legal methods:

"[…] Be that as it may, for the time being it is not we who are being destroyed by legality. It is working so well for us that we would be mad to spurn it as long as the situation lasts. It remains to be seen whether it will be the bourgeois and their government who will be the first to turn their back on the law in order to crush us by violence. That is what we shall be waiting for. You shoot first, messieurs les bourgeois.

No doubt they will be the first ones to fire. One fine day the German bourgeois and their government, tired of standing with their arms folded, witnessing the ever increasing advances of socialism, will resort to illegality and violence."

The egalitarian past (and future?) of politics

Let's say that politics is the formation and use of coalitions for determining matters of social concern within a group.

I was listening to a talk about human evolution as a process of finding a cognitive niche by Andrew Whiten when he made two critically important points which, while I was aware of them abstractly, I'd never previously considered as points about the inherent character of politics.

1. Among early hunter gatherers, we have strong evidence to believe that political coalitions were far more likely to be used to *oppose* the ascension of a person to a position of social dominance than support it. This is well known- there's a tonne of research on counter-dominance strategies among hunter gatherers, and the point that most early hunter gatherer groups were politically egalitarian in addition to economic egalitarianism is well understood.
2. Even among chimpanzees, some of the primary uses of coalitions include supporting the interest of groups of weaker individuals against a single stronger individual, and acting as king-makers, preventing a single male exercising absolute dominance.

It occurred to me that these days we generally view politics as a tool of the powerful, and yes, they are. In the very beginning though, we have reason to believe that politics was an innovation of those individually less powerful- the formation of coalitions to control charismatic or physically powerful individuals.

In other words, pervasive use of politics, was once a feature that separated us from certain other types of animals, making us far more egalitarian than, for example, chimps. Almost unimaginably, the very feature of our behaviour which made early societies so much more egalitarian than many other primates would one day enable certain individuals to possess billions of times more wealth than others.

Our challenge then is not just to explain why human societies have moved from the egalitarian to the inegalitarian (and in some cases, part of the way back again), but to explain why and how the primary use of politics was transformed from counter-power to power. How did the orientation of politics become inverted from its original use?

Reevaluating our conception of the origins of politics might enable us, as people concerned with a more egalitarian world, to relate to politics differently, and with less revulsion, viewing it most naturally as 'our' tool not theirs.

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

The alternative conception of politics has, after all, never quite gone away. Even today, leaders of human groups are required to at least *pretend* to rule in the common interest, a throwback inherent in the way political power is structured, as an exercise in the perceived interests of the many who form a coalition.

Trying to envisage a road to equity is a fearful task, but we shouldn't be too afraid of that most maligned of tools, politics- ancient friend of the many, foe of the few.

PART 7: POETRY

Deadwater

I

I recall in tranquillity

Fever-dive hours.

Once I saw a sailboat listing

Upon a great-waved sea

The sea was I and so was the boat

I could not see any stars

For the blasts of ocean-spray

In what quiet cove can I go hiding from a storm

Blasting up the cartoid artery and flooding through

The cognitive estuaries, over-spilling memory's tributaries?

Tell me where I might make my stand against my wrath?

Might a clever present play the future off against the past?

Am I to live only in the lacunae between foretelling & recollection

In the times between guilt and dread when, exhausted of mental flight,

Whether backwards or forwards, the I drifts in easy content?

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

We shall build a tower

let us make us a name, lest we be scattered abroad upon the face of the whole earth

II

Behold, a shattered glass bowl that held doubts

They multiply in shattering

As each beam of light

Crosses every glass splinter

It breeds a new splinter

And a new lance of light

Fecund heresiarch

Absolute clarity lies within

That lit glass rubble but the trouble

Is that so does everything else

As in Borges' library up in that tower

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

III

Do you know where your right hand is? Walking through a shop and not knowing whether you've assaulted someone heedlessly. Analysing each moment of your past like a sicko prosecutor. The fears iterate by sinister Darwinism, seeking cognitive blind-spots. Did I mutter threats of violence to that child? Did I insult that shop attendant? Mixed memory and aversion form a rancid bin-juice born decaying.

IV

I came to the stairs

There was a wobble in her voice

By each step her voice rose higher

So I rise to her and she calls with greater urgency

And I rise to her with greater urgency

She and I can only meet after escalation shatters

Past the horizon of panic and further-

Past the sea rock of worn defeat

She and I must be one.

I sprint.

V

Imagine that someone came to you in the middle of the night, stepped into your mouth and began to grow through your capillaries. They were not content merely with habitation, their constant insistence was that you must keep grafting dead organs and limbs onto yourself. You become a born-again Frankenstein (don't be a pedant) with all the zeal of a convert to an undead lifestyle. The new limbs are heavy, and stink, and burn up your flesh with septisemic fire and puss-flood, but the man who stepped inside your mouth begs you stitch on more.

VI

The inside of a head becomes lonely as it becomes crowded

The only things that elbowed through those crowds

Were other hauntings

Brief dune-sedge love in salted ground

Warring wrath against money made world

Twin engines of raging-love and loving-rage

Racing for diversion and the exaltation of rebellious motion

Circulation round the track kept my blood in motion

Rammed down winds to bellow my lungs

Political contention, war, courtship, frenetic study

Vain dreams of greatness, discontent

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

Which gave me a little contentedness

To declare permanent war or endless love

And so to terminate surrender in unutterable resolution

“Optimism of the will!”—clenched hands, though they wobble

In the obsidian lands where resistance gave no comfort

Resistance still gave sustenance

Just as all the previous Sugatas

VII

Life is so long. Are you so innocent? You are tired. You dream of a gentle place. You saw it as anyone might imagine it—holy light on wild-flowers, easy with its comforts, free with its joys. To be such a place it had to be distant from this world and sealed against you.

VIII

Maybe I just wasn’t fucking often enough?

Victorian life is better novelised than lived

Hysterical, neurotic, guilty, phantasmal

Maybe I wasn’t drinking enough?

A friend called me the Ayatollah

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

In respect of my beard and sobriety

Hume and the Buddhist sages pronounced that persons are aggregates without greater unity. I find myself a bundle but there is no liberation here. The parts rub against each other like cans in a grocery bag bruise fruit. Or perhaps I am the curate's egg.

IX

Give me a seabird's wings

On the cliffs, about forty meters over the crab pools

I dream of ascending with the gulls, but higher

Diving and again rising in alliance with wind

What waves perturb the gull are brief

And if it is to end by hawk, that too is brief

Yet I would rise higher still, till I sat on a perch

Overlooking time and the jolting succession of moments

Above the waves of kings, ministers, exchequers

Yet if I am not to reach that exalted perch

I will be low enough to observe the bright net

Of refracted sun that plays upon the hills of water

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

Give me a seabird's wings

X

Easier perhaps to talk of the accoutrements of terror and the reflections it invoked. Easier to do that then to photograph medusa. Yet I do remember being confused as to whether I was more guilty or more afraid. It seemed important that I be more guilty than be afraid, but it is hard to feel guilt while facing knives. Consequently, I felt supplementary guilt at my thin guilt.

We shall build a tower

let us make us a name, lest we be scattered abroad upon the face of the whole earth

XI

The future is boundless, not only ahead but sideways

The patterns of your inferences only ever ape

The subtle causal chains which bind the forward momentum

Of the world whose surface you cling to

The mind is stretched between times and possibilities,

Beyond any accommodation by mental sinew and bone

The heart successively roars and fizzles

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

XII

I came to the living room

And it was filled with ash

Though I never smoked

Or sat by fire

I made an ink of that ash

And began to write these verses

upon my arm

XIII

He is there, and I smile into his oblivion

He never loved you, so ideas of romance

Had the character of Banach-Tarski's sphere

He is gone now, other suburbs, other worlds

I do not miss him, except on special occasions

My affections were never lost, except perhaps at the first moment

Dead on arrival

Yet still worthwhile

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

It is right to rebel against most things

But not you, oh sweet tyrant

It's good odds you kept me breathing

IXV

We do not sit upon heaven's throne

Nor are we the rebel, cast down like a slash of lightning

We are the flesh that raised our gaze

Half wondering, half begging

The dance is ending, where is the bridgegroom?

XV

How rash are those who clamour for justice?

(I have been among them)

Life is wide, deep and changing. We are excesses

Of identity, act, motivation.

Of miscalibrated judgement and selfish grasping.

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

Do you think you would be clean under heaven's eye?

Were there a book that contained each numbered thought and small deed

Of yours wouldn't you shred it, burn it and eat the ashes?

I wouldn't. I would give you that book. Press you to read it.

I do not think you would like me, but my terror is to be misunderstood

I fear that you will think I am a different kind of monster than that I am.

So I give you my promise, that should an angel scribe that book

I'll give you a copy.

And I promise that if you ever give me a copy of your celestial biography

I'll try to shut the my eye of judgement and open that of mercy

It's simple self interest. Chesed pro chesed.

XVI

Can we remember pain? In our mind's eye we might

See rose fluids or, under that, a startling glimpse of pearly white

Laid open by a scalpel. We shudder back. We peer forward.

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

But who has the pen by which to bind agony?

“Sharp”, “dull”, “throbbing”, “irritating”, “intense”

Wholly feeble, as if a snake tried to wander with its vestigial leg bones

But that is where we find ourselves—thirsty for conveyance in a desert of names

We can only hope to articulate pain through our inarticulateness

Just as, by chance, static on a television set captures a snowstorm

I remember wandering the streets, sobbing and calling for divine fire to kill me and all the other wicked. As I wept I listened to pop on half smashed headphones. What would it take to make you march through city streets weeping and calling the fires of an unknown God?

XVII

I ascended to the attic

To store, retrieve, invent

A mnemonic parade

Without volition my hands

Raise the dust in small incantations

How does one dislodge a fake memory?

Or terminate the routine of shuddering

I see

He and she are here, interlocked eye-beams

I am not in either eye

In this attic I lay in the pattern of my veins

I am sinews. Whether these gobbets

Be thought or flesh I am in neitherway free

I am chained by my own substance

Above me powers contend in the air.

XVIII

Think now

Life has many cunning passages, contrived corridors

And issues, deceives with whispering trepidations,

Guides us by vanities.

After such knowledge what forgiveness?

Forgiveness after such knowledge what?

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

What forgiveness after such knowledge?

Knowledge what forgiveness after such?

Such knowledge what forgiveness after?

IXX

In metamorphosis the tissue is not merely subtracted from and added to inside the pupae, rather the whole flesh devours itself, save for microscopic clusters (imaginal bodies), becoming a soup of cells. What unites both life-stages is scarcely more than a double-helixed teleos. Yet memory persists.

We shall build a tower

let us make us a name, lest we be scattered abroad upon the face of the whole earth

XX

If I could but seize the wax of Icarus

The tailor of Ulm's fabrics

Etana or Bladud's crown of feathers

If I could but fly, I could seize the sun's silver

Forge a mirror by which to demonstrate

The storm that rends the head

Of some shivering soul you know

Forgive a thief that stole for you and

Shelter all, for you cannot see their weather

XXI

To find a point of collapse at which

loss and victory die.

And that sea is now

A vast lake that

Night or day

Forms a perfect twin

To the sky

Over the stones of the tower

Drift currents and sweet, lazy fish

The waves will dance again

But I might hope to dance

With them

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear

Afterword

A word on credit. This poem is allusive to the point of plagiarism, and past that (about 5% is lifted from other poems). My purpose is to convey an experience with all that I have and I'll gladly steal words for that. I have no concern to prove myself as a poet, only to tell the story as well as it can be told.

The debt to T.S. Eliot is obvious, even in the title. The debt to the Aiken's Tetelestai and the Romantics (including Eliot perversely read as a romantic) is less obvious. It's very much a poem about me, and I apologise for that vanity. My story is not unique. My particular kind of OCD based on a fear of harming others is quite common. Yet few talk about it for fear of seeming like a dangerous weirdo. It is an inherently self-concealing form of mental illness. Especially as I've gotten older, I've tried to avoid the narcissism of self-display even in an anonymous form, but I want to show you this story, lest it be scattered everywhere among the nameless like me, and forgotten.

For those who have loved me.

You might also enjoy my blog: https://philosophybear.substack.com/ and my subreddit: r/philosophybear