

Book Review Runners Up

Volume 1. A -R

[A History of Generation Ships in Science-Fiction](#)

[A Libertarian Walks Into a Bear: The Utopian Plot to Liberate an American Town \(And Some Bears\) by Matthew Hongoltz-Hetling](#)

[An Evolutionary Theory of Economic Change by Richard R. Nelson and Sidney G. Winter](#)
[Habit and Routine, Not Calculation and Optimization](#)
[Economics Has Come A Long Way](#)

[Antifragile: Things That Gain from Disorder by Nassim Nicholas Taleb](#)

[I. Prologue](#)

[II. The Triad](#)

[III. Barbell Everything](#)

[IV. Killing Volatility](#)

[V. Against the Nerds](#)

[VI. Against the Predictors](#)

[VII. Via Negativa](#)

[VIII. Against the Immortals](#)

[IX. Autobiography](#)

[Anna Karenina by Leo Tolstoy](#)

[Apocalypse Never by Michael Shellenberger](#)

[I](#)

[II](#)

[III](#)

[IVa](#)

[IVb](#)

[V](#)

[Astounding: John W. Campbell, Isaac Asimov, Robert A. Heinlein, L. Ron Hubbard, and the Golden Age of Science Fiction by Alec Nevala-Lee](#)

[Atlas Shrugged by Ayn Rand](#)

[Quick reviews from 1957](#)

[Characters](#)

[Review](#)

[Blind Spots](#)

[Closing remarks](#)

[Atlas Shrugged by Ayn Rand](#)

[Bad Religion: How We Became a Nation of Heretics by Ross Douthat](#)

[The Heresy of Christian Nationalism](#)

[The Heresy of “Pray and Grow Rich”](#)

[The Heresy of the God Within](#)

[The Heresy of “Brave Certainty”](#)

[Why is Christianity in decline?](#)

[Who are the “Nones?”](#)

[What is Douthat’s Defense of Orthodox Christianity?](#)

[But let’s talk about “Accommodation.”](#)

[Can Christianity Recover?](#)

[But how should Christianity respond?](#)

[And What About the Nones?](#)

[Behave by Robert Sapolsky](#)

[Everything is connected](#)

[Biology is too hard](#)

[Brave New World by Aldous Huxley, Or, Should I wirehead my cat?](#)

[Capital as Power: A Study of Order and Creorder by Jonathan Nitzan and Shimshon Bichler](#)

[Part A: Overview](#)

[1. A Need for Better Theory](#)

[2. Business and Industry](#)

[3. Accumulation of Power](#)

[Summary of Part A](#)

[Part B: Dilemmas of the Dismal Science](#)

[1. Politics and Economy](#)

[2. The Nominal and The Material](#)

[3. What is Value?](#)

[Summary of Part B](#)

[Part C: The Machine](#)

[1. Internal and External Logic](#)

[2. Capital and Capitalization](#)

[3. Profit and Differential Profit](#)

[Summary of Part C](#)

[Part D: The Machine in Context](#)

[1. Economic Implications](#)

[2. The Space of Political Systems](#)

[3. Connections with Other Theories](#)

[Summary of Part D](#)

[Should You Read This Book?](#)

[Che Guevara: A Revolutionary Life by Jon Lee Anderson](#)

[Consciousness Explained by Daniel Dennett](#)

[Introduction](#)

[The Cartesian Theater](#)

[What Makes Consciousness So Confusing?](#)

[The Cartesian Theater](#)

[The Multiple Drafts Model](#)

[Do We Mean What We Say?](#)

[Awareness](#)

[The Unconscious / Conscious Boundary](#)

[The Mind's Eye](#)

[Qualia](#)

[Philosophical Zombies](#)

[Conclusion](#)

[Creative Evolution by Henri Bergson \(1907\)](#)

[Ch. I](#)

[Ch. II](#)

[Ch. III & IV](#)

[Other Quotes](#)

[My Comments](#)

[Defenders of the Truth: The Sociobiology Debate by Ullica Segerstråle](#)

[Wilson and Lewontin at the center](#)

[At the centers of different worlds](#)

[Planters, weeders, couplers and decouplers](#)

[Dawkins and Gould: the next generation](#)

[Modern day descendants](#)

[Circling back](#)

[Democracy in America vol. I by Alexis de Tocqueville](#)

[Disunited Nations \(2020\) vs Dawn of Eurasia \(2017\)](#)

[The US World Order is coming to an end](#)

[The US Navy patrols the world](#)

[Zeihan's geographic determinism](#)

[The Dawn of Eurasia](#)

[Determinism vs Hope](#)

[How much of the future can anyone ever see?](#)

[Reversion vs Evolution](#)

[The Russian soul](#)

[Lighting the path ahead](#)

[Driven to Distraction: Recognizing and Coping with Attention Deficit Disorder from Childhood Through Adulthood by Edward M. Hallowell & John J. Ratey](#)

[Hypothetical Condition](#)

[Case Studies](#)

[Other People](#)

[Treatment](#)

[Conclusion](#)

[Down and Out in Paris and London by George Orwell](#)

[Enlightenment Now by Stephen Pinker](#)

[Exploring Reality: The Intertwining of Science and Religion by John Polkinghorne](#)

[Reality](#)

[The Causal Nexus of the World](#)

[Human Nature: The Evolutionary Context](#)

[The Historical Jesus](#)

[Divine Reality: The Trinity](#)

[The Nature of Time: Unfolding Story](#)

[The Spirit and the Faiths](#)

[Evil](#)

[Ethical Exploration: Genetics](#)

[Imaginative Postscript: Some Naive Speculations](#)

[Conclusion](#)

[Fantasyland](#)

[I.](#)

[II. A summary-summary](#)

[Part I: The Conjuring of America: 1517–1789](#)

[Part II: United States of Amazing: The 1800s](#)

[Part III: A Long Arc Bending Toward Reason: 1900–1960](#)

[Part IV: Big Bang: The 1960s and '70s](#)

[Part V: Fantasyland Scales: From the 1980s Through the Turn of the Century](#)

[Part VI: The Problem with Fantasyland: From the 1980s to the Present and Beyond](#)

[III. The Good](#)

[1. The hidden connection](#)

[2. Reframing culture](#)

[An Aside: Boy Scouts](#)

[3. Combining real and fake](#)

[4. Americans believe crazy things](#)

[IV. The Bad](#)

[1. Lack of Explanation](#)

[2. Out-of-hand dismissals and misrepresentations](#)

[3. Excessive credulity](#)

V. Conclusion

Farmers for Forty Centuries by F. H. King

Feed by M.T. Anderson

Fitzpatrick's War by Theodore Judson

Games People Play: The Psychology of Human Relationships by Eric Berne

Background

Transactional Analysis

Games

Game - Examples

I'm only trying to help

See what you made me do

Vindication

See what you made me do

If it weren't for you

Debtor

I'm only trying to help you

Look how hard I've tried

Self-reprimanding, obtaining forgiveness

Alcoholic

Schlemiel

Confirming worldview and stereotypes

Kick Me / Why does this always happen to me?

Debtor

Reassurance

Courtroom

If it weren't for you

Why don't you - Yes but

Feeling power & Control over others

Now I've got you, you son of a bitch

Kiss Off

Interesting observations

Complementary games

Reversible roles

Rough model

Children are impressionable

Sabotage

Antithesis: How to Stop Games

See What You Made Me Do - Vindication

Look How Hard I'm Trying - Vindication

[Self-reprimanding & Obtaining Forgiveness](#)

[Why don't you ... Yes but - Reassurance](#)

[Control and Power over others](#)

[Epilogue](#)

[Gödel, Escher, Bach by Douglas R. Hofstadter](#)

[I. Formal systems and interpretations](#)

[II. Jumping out of the system](#)

[III. Truth vs. provability](#)

[V. Self-reference and the proof of Gödel's theorem](#)

[VI. GEB Part II](#)

[Guns, Germs and Steel: The Fates of Human Societies by Jared Diamond](#)

[How Children Fail by John Holt](#)

[Human Scale Revisited by Kirkpatrick Sale](#)

[I](#)

[II](#)

[III](#)

[IV](#)

[Human Smoke by Nicholson Baker](#)

[Aerial Bombing](#)

[Churchill](#)

[On blockades:](#)

[Churchill on the Thickness of Shells:](#)

[Roosevelt](#)

[Roosevelt on Entering the War:](#)

[On Entering the War:](#)

[Churchill on Hearing about Pearl Harbor:](#)

[Pacifists](#)

[Gandhi on Pacifism:](#)

[Nazis](#)

[On Kristallnacht:](#)

[On Sterilization:](#)

[On Gassing:](#)

[On Mass Graves:](#)

[European Jews](#)

[Anti-Semitism In Lithuania:](#)

[Churchill on Jewish Foreign Nationals:](#)

[Churchill on Palestinian Jews:](#)

[Roosevelt on Immigration Quotas:](#)

[Roosevelt on Relocation:](#)

[Roosevelt on the Wagner-Rogers Child Refugee Bill:
Why You Should Read This Book](#)

[On the Title:](#)

[Last Tsar. Nicholas II, His Reign & His Russia by Sergei Oldenburg](#)

[Loserthink by Scott Adams](#)

[Think like a Psychologist](#)

[Think like an Artist](#)

[Think like a Historian](#)

[Engineer](#)

[Leader](#)

[Scientist](#)

[Entrepreneur](#)

[Economist](#)

[Comedian](#)

[Using the Ego as a Tool](#)

[Conclusion](#)

[Made to Stick by Chip Heath and Dan Heath](#)

[1.](#)

[2.](#)

[3.](#)

[4.](#)

[5.](#)

[6.](#)

[7.](#)

[8.](#)

[9.](#)

[On Numbers and Games by John H. Conway](#)

[A Divine Game](#)

[Human Games](#)

[Numbers](#)

[Worlds](#)

[Open Borders: The Science and Ethics of Immigration by Bryan Caplan and Zach Weinersmith](#)

[The argument](#)

[The objections](#)

[1: Parochialism](#)

[2: Humility](#)

[3: Environment](#)

[4: Culture](#)

[5: Inequality](#)

[Conclusion](#)

[Appendix: Other Arguments I Would Have Used](#)

[Power, Sex, Suicide: Mitochondria and the meaning of life by Nick Lane](#)

[I - What is this book about?](#)

[II - What is this book actually about?](#)

[III - But now, really, what is this book actually about?](#)

[IV - So, is that all?](#)

[Probability Theory: The Logic of Science by E. T. Jaynes](#)

[On the origin of PT:TLoS.](#)

[What Jaynes taught.](#)

[À la recherche du temps perdu.](#)

[Paradox lost.](#)

[Exegi monumentum.](#)

[Reasons and Persons by Derek Parfit](#)

[Reinforcement Learning: An Introduction by Richard S. Sutton and Andrew G. Barto](#)

[What is reinforcement learning](#)

[How reinforcement learning works](#)

[Mistakes drive learning](#)

[Dopamine and Happiness](#)

[Simulated Experience](#)

[Simulated Rewards](#)

[Importance Sampling](#)

[Conclusion](#)

[Revolt of the Public by Martin Gurri](#)

A History of Generation Ships in Science-Fiction

(This essay contains multiple spoilers about several science fiction novels; I will put the spoilers all in footnotes at the bottom: SO DON'T READ THE FOOTNOTES IF YOU WANT TO AVOID THE SPOILERS)

In 1958, British World War II veteran Brian Aldiss published his first novel, one that relied on some of his experiences during the Burma campaign in the humid Southeast Asian jungle.

In the novel, a group of jungle dwellers wanders around their strange environment, and the reader quickly realizes that the jungle has grown inside a gigantic starship. Multiple adventures occur before the protagonists understand that they live inside a sort of interstellar ark in which their ancestors lived, died and bred for generations. Eventually, something even more astounding about their predicament is revealed (1).

Aldiss' novel, called "Non-Stop," was the founding piece for a new kind of science-fiction tale: stories about generation starships, a hypothetical type of craft that travels at sub-light speed while its human occupants spend generations inside, waiting to set foot on another world.

These tales are interesting in themselves, and as illustrations of current concerns, as well as plans and hopes for the future expansion of the human race. Some generation ship novels are among the best sci-fi books ever written; many others are frustrating pieces overflowing with ideas but also prejudices, clichés and particular obsessions that reflect our age only too well. When we discuss generation ship stories, we not only discuss the future of an idea, or the present state of that idea: we also discuss what writers have imprinted in their minds when they execute that idea.

"Non-Stop" may not have been the first actual story to depict a generation ship, but is definitely the first complete example of the genre, since it contains almost all of the features later to be seen in similar novels: characters confounded by their situation; characters who come to forget the reason why they're floating in space in the first time, or doubt that they're actually floating in space; internal struggles for power that go beyond anything the mission designers anticipated; frustration; sickness; swift death; a melancholy reflection on the fragility of human societies.

Even conceding that fact, there are the many varieties of generation ship narratives out there: there are, for example, multiple stories with hibernated crews that are unfrozen at the wrong time, or too early, or too late, or are killed while frozen, or are frozen and defrozen at specific points by a childish AI ("the Chimp") leading a crazy mission spanning billions of years, such as in [Peter Watts' outstanding Sunflower Cycle](#).

Such narratives, while often similar to those in actual generation ships, are fundamentally different in that it's the very same people who start and end the trip. In fact, the defining characteristic of the generation ship tale is that "generations" span

between the start and the end of the journey: at least two. Otherwise it's just a very long trip, with little to do.

A good, little-known example of the non-generational ship tale is "[The woman from the ocean](#)," a short story written by Karl Bunker, published in 2014. It's about an astronaut who comes back to Earth far in the future, with news that colonization of space is impossible because everywhere else in the universe is just too inhospitable.

Like the dudes from the Planet of the Apes, she returns to a much-changed Earth of the far future, due to relativistic effects: society there has regressed to agriculturalism for interesting reasons. In a way, "The woman..." is a sort of variation on "Non-stop": even if the travelers stay sane and focused on the mission, it may be our whole planet that goes crazy and primitive and back to, literally, the law of the jungle.

The interesting thing about "The woman..." is that the protagonist has been to other star systems, and she has returned to Earth in defeat, the last survivor of a doomed enterprise which message is: Earth is where you should be, don't have silly dreams. This is surprisingly often the message of generation ship stories.

Take [the 2015 sci-fi novel "Aurora"](#) by Kim Stanley Robinson. Like "The woman..." it has a strong independent woman leading an ambitious missions to disappointing, but oddly comforting, results (2). Both stories underline the harshness of the universe out there and how Earth can be the only place where humans can ever hope to survive, a theme with evident environmentalist undertones.

When "Aurora" begins, the ship is starting to fall apart after almost 200 years out in space, just as it's reaching its destination in Tau Ceti (3). Devi, a 5th or 6th generation traveler stuck in the ship, is a woman genius who fixes all sorts of problems arising, and in her spare time teaches the AI how to think proactively and narrate its story.

Devi has a husband who is the personification of blandness, Badim, who somehow survives the entire novel without saying or doing anything memorable even though he's in the thick of everything. This pair has Freya, a daughter that Devi herself describes as a bit dumb, an example of how the "zoo effect" is making the humans aboard Ship (it has no name, strangely, this Ship) dumber; but Freya confounds her mother's expectations and eventually becomes the mission's leader. (4).

A very relevant fact is that the Ship, as depicted by Robinson, is very much in line with mainstream SF conventions of how future colonization should work, and is perhaps the classic standard-issue Generation Ship: it has plenty of animal and human diversity, but the differences between humans stop at the skin-color level; underneath, they are all the same, enlightened modern person of the future type with just the right amount of funny foreign names: religion, ethnicity, race, nationality and language differences have all disappeared, and all there's left is some holidays (we have Halloween, you have Día de los Muertos!) and food specialties.

Really, Robinson should teach the world how he managed to squeeze Jews and Muslims, Russians and Poles, Koreans and Japanese in the same ship for generations, with no problem at all! His generation ship is, much like Robinson's Mars, an extension of Northern California without the unsavory bits. There are no pious Muslims objecting to pigs running around in the ship, no Spaniards objecting to the objection on the

grounds that one can't spend 200 years without ham. Latinos happily speak inglés, sí, señor, at all times with clear crystalline accents, so Freya can understand.

Like many other writers, Robinson is a firm believer in "The Lathe of Heaven principle," the notion that in the future we'll have every current social problem except racism, and racism is absent from his mission (5). Another beloved, common characteristic of generation ships is the addition of areas with varied landscapes and climates, mimicking Earth, but often with no sense or purpose other than wasting space with pointless decoration. Whoever designed Robinson's Ship decided that having a Mongolian biome and Mongolians was really important; but there's no Chinese biome, even though there are 300 times more Chinese than Mongolians on Earth, approximately.

In "Aurora," the trip is OK and things really go downhill for everyone involved when it ends. The point is: even if the colonists survive the journey, bad, bad things can (and will!) happen at the destination (6). In "[Ark](#)" (2009), one of many excellent novels by Stephen Baxter, the faster trip (just three generations for this generation-ship story) is truly hellish, since the colonists were forced to leave Earth just before the planet was made inhabitable, and they have the extra strain of knowing they are the last hope for the human race.

Colonists in "Ark" fight, rape and kill each other while stranded in two narrow hulls with no fake Mongolian steppes to excite the mind. Grace Gray, the strong independent woman who is Ark's protagonist (and who eventually passes on the torch to her daughter Helen, also strong, and independent), is a "gate crasher," one of a handful of people who were not supposed to be saved in the Ark but gate-crashed her way into the ship anyway, which doesn't really improve moral within the mission.

On arrival in the planet they call "Earth II" in 82 Eridani, things are not so rosy as they anticipated (surprise!) The planet is barely livable at all, and Grace, like the hopeless Freya, takes a pretty radical decision (7). This decision, to Baxter's credit, somehow makes a lot more sense in "Ark" than in "Aurora."

A nice touch in "Ark" is that, again unlike the case with "Aurora," the narration keeps in touch with the dissidents from Grace's and Freya's motherly regimes. The second hull travels all the way to Earth III, and things get dicey when Wilson, a daredevil pilot who becomes the leader of a rapey, violent gang -- the usual outcome for strong independent men in Generation Ship stories -- convinces his idiotic men-mates that they're not actually floating in space but living in a virtual reality (8). From Aldiss' times, the idea that people inside of space tin cans will go crazy in such ways is widespread.

The Wilson archetype often finds his comeuppance at the hand of strong independent women in stories involving Generation Ships and interstellar colonization. A perfect example comes from another Baxter novel, "Proxima" (2013), in which a bunch of desperadoes are dumped on a rocky planet around Proxima Centauri that has breathable air, as a sort of Australian-style colonization plan.

Mardina, a strong independent woman who was one of the female guards protecting the desperadoes in ship that took them to Proxima, is left behind with them and she eventually comes across the gang leader of a "Mad Max"-style group of survivors: instead of discussing his crimes, she quickly grabs an arrow from her quiver and,

Hunger Games-style, shoots this Wilson dead, ending his short, bloody reign of terror and releasing his henchmen from terrified obedience.

Wilson-like dudes have an even harder time of it in "[Seveneves](#)," [Neil Stephenson's 2015 classic](#), in which the unexplained disintegration of the Moon leads to the slow destruction of our planet via bombardment and the hurried construction of a sort of Generation Ship/Space Station in orbit around Earth. The Seven Eves of the title are the seven (!) strong independent women who eventually survive the ordeal, after every single man, including many Wilsons, kill each other (9).

Stephenson has made a career of writing about strong independent women in spaceships who won't take either a yer or no for an answer, or really any answer at all, from cowardly stupid men. In his long, thrilling, Generation Ship-like "[Pushing Ice](#)" (2005), a wild ride into alien technology and manufactured universes, it's two women rather than one who strongly, independently push mankind towards its furthest reaches.

The predominance of female protagonists in sci-fi is, of course, a novelty of recent decades. In the supposedly golden era of sci-fi when Aldiss started to write, men led the struggles along the way, and those on arrival, with most female protagonists as love interests/maidens in despair. But one thing remains the same: mankind's worst instincts often express themselves in full inside the hulls.

Honestly, it's no wonder that many people inside Generation Ships just prefer to forget about the mission, and enjoy however little is enjoyable inside the steppe-less ships. In [Richard Russo's "Ship of Fools" \(2001\)](#), nobody in the spaceship Argonos can really tell why they're traveling through space, where they're heading and what they will do when their descendants arrive. In fact, no one on board can even say for certain where the ship came from (10).

At least, in "Ship of Fools" people have a certain idea that they're something called human, and they organize themselves along societal lines that any human can recognize: a class-system has developed in which the commoners and poor serve in the lower levels and provide maintenance for the ship, while the aristocrats rule and recommend that the poor eat cake, whenever they complain about their lot.

In [Greg Bear's "Hull Zero Three" \(2010\)](#), forgetfulness reaches a whole new level: one spends dozens of pages with a man who wakes up from a dream-like state, naked and freezing, with no memory. He's led by a little girl, who calls him Teacher, through a series of corridors in a generation ship in search of warmth. Everything is broken down and malfunctioning and the huge craft is full of strange non-human creatures, who join the Teacher and the Girl in a diversity fan-fest/quest to understand just what the hell is going on.

Eventually, the reader comes to the realization that Teacher is a clone who has been resurrected several times before. Clones are a nice, relatively recent addition to the Generation Ship genre, and one that is best, and most forcefully, developed in "[Noumenon](#)," a 2017 novel by Marina J. Lostetter.

*

"Noumenon's" central conceit is that generation ships would be best manned by clones re-cycling copies of themselves.

As the novel starts, young scientist Reginald Straifer has discovered that the star LQ Pix may, or may not be, surrounded by a Dyson Sphere; this is the late 21st century, and the world has been thoroughly Kim Stanley Robinson-ed: decent, multi-ethnic technocrats rule, conflict has diminished, and there's plenty of money to set up several long-range missions, no questions asked -- so, to his surprise, the young scientist manages to get a huge amount of resources to send a ship all the way to the alleged Sphere.

After careful consideration, Straifer decides that he prefers to send an entire fleet of mid-size ships, connected through shuttles, rather than a large mothership. This will travel a huge distance using some sort of modern technology that allows flying at significant fractions of the speed of light; and then, for some reason not really entirely explained, return to Earth to report on the findings instead of just mailing them.

Since the destination is very far and the trip very long, even at huge speeds, a Generation Fleet is constructed with multiple ships devoted to specific tasks -- one for biomes in which to live the illusion that one is on Earth, some for scientific research, others for manufacturing or living quarters -- and one ship shines above all: a reproduction ship, in which genetic copies of all of the fleet members are stored, so that clones can be produced at will.

Reasons are given for the preference for a clone crew but, honestly, they are not very convincing. The main reason is that the crew must be very large to account for attrition, attend to the mission needs, etc, and Earth can't really afford to lose so many brilliant minds. Really? That means a bunch of brilliant, capable people are selected, and then cloned: their clones are raised in a sort of massive hippie-techno commune in Iceland, because our obsession with Nordic minimalism hasn't gone away by the early 22nd century, apparently; then, at the age of 20-something, the mission is sent off into space.

I wouldn't vote for this arrangement for that sort of trip. But this is the arrangement. And it leads to a series of consequences, only some of which are predictable.

For example, people form couples, but can't have children: they receive clones to raise in the ship; they themselves are "put out of action" at a certain age, so that the crew won't be too elderly, but a younger clone of themselves, who they may know very well, will be around and be, for example, David Román II, followed by a David Román III and a David Román IV, and they can coincide for a while, and eye each other warily (10).

There's not a lot of voting here. The position of mission captain is actually inherited by the clones of the first captain, and the same happens for several key positions such as those of head of communications, research, etc.

This arrangement is pretty unique in the history of sci-fi (11), and leads to all sort of interesting plot lines (12). It really is striking to realize that things have been ordered so that the same social relationships that were set up at launch in the 22nd century are to be replicated ad infinitum, for the duration of the mission.

There's a Latino clone, Diego Santibar, who literally spends several generations as a ship gardener (a high-end gardener, though!) At one point in the novel, it seems that one of the Santibar iterations is kind of tired of being around plants and one breathes a sigh of relief: could it be that there's a crack in this immense façade of genetic

determinism? No. It turns out that Diego's supposed clone is actually somebody else's clone: thus his dissatisfaction with Diego's eternal lot. The understanding is that Santibar's next actual clone will be very happy with a return to his gardening.

Something similar happens with John Mahler, the aforementioned mission captain, who remains the same mission-obsessed creep with authoritarian tendencies throughout, despite a colorful change of jobs (13). Margarita, the head of communications (with Earth), remains exactly the same corporate drone through centuries of iterations. The clones of Sailuk, the Eskimo wife of captain Mahler, keep marrying Mahler clones, as if proving a theory of love or attachment imprinted in the genes.

The biggest element of social and mission flexibility is provided by the most stable character in the mission, the Artificial Intelligence that helps the crew control the ships and everything that happens in them. This AI is supposedly based on a cellphone AI that Straifer I used to love, and is very well described: it is in fact, one of the most credible characters in the novel; an artificial sentient being that sticks to the mission but has leeway to adjust and shift priorities as long as they don't conflict with the ultimate, set-in-stone goals: a milder, better-mannered version of Watts' Chimp.

It really is endearing how Lostetter and her characters remain faithful to the book's hard notion of genetic determinism. But the book's final pages, in a very distant future with very weird humans, don't really work as well as the precedent parts. One big lesson of Generation Ship stories is that we 21st century humans may not necessarily understand or be interested in the doings and feelings of our very distant descendants, who by necessity just don't have a lot in common with us -- that is, unless they are clones of the same 21st century people, obsessively reborn.

In "Noumenon," the clones are so much more sympathetic than 4th millennium humans. In "Sevneves," the triumphant descendants of the epic survivors are just entirely uninteresting, being so very detached from our reality; much the same happens with the colonies left behind by Baxter's "Ark," even after a much shorter time span. In Watts' Sunflower Cycle, shipbound humans are terrified of even casting a glance at the things that will descend from humans and live millions of years from now.

Our connection with our far-flung descendants in the stars, or even low-Earth orbit, will one day become very, very thin. We can only hope that these descendants still think of themselves as humans, and think of us as worthy ancestors, rather than hopeless brutes dominated by clichés and group-think, slaves to literary fashion: perhaps, there's a couple of things we can do to help our cause, and prove we were indeed worthy.

SPOILERS BELOW:

1. That the starship is actually orbiting Earth, having gone to another star system and then returned when the human colonists encountered an alien pathogen that almost wiped them out and caused the societal collapse in which the protagonists live early in the book.
2. In both books, the leading characters are happy with having managed failure rather well, by the low standards they set for themselves.
3. An illustrious place in SF history, where Downbelow Station is located. That

fictitious station gives the classic 1981 novel, written by the actual strong independent woman Carolyn Janice (C.J.) Cherry, its title.

4. Robinson has stated that Freya is his avatar in the novel, which isn't an entirely flattering statement. It must be said Robinson's leading heroes in his famous Red Mars trilogy were also pretty frustrating people to spend time with.
5. In "The Lathe of Heaven," author Ursula Le Guin presents a future in which all humans are race-less and gray. To my knowledge, there's only one generation ship novel in which racism plays any role, "[An unkindness of ghosts](#)"(2017), a straightforward, extended analogy for American chattel slavery.
6. A prion kills the explorers left on the moon that the Aurora mission planned to colonize. Freya then becomes the spokeswoman for history's most defeatist movement: she gets half of the people in the Ship to vote for cutting and running back home. It's important to underline that, in "Non-stop", there is no clear explanation of the pandemic that led to the ship returning Earth, so it's all vague and easier to take: in Aurora, there's enough detail to infuriate those of us with an advanced grasp of virology.
7. She leads a faction of the colonists in one the ship's hulls back to Earth, which turns out to be yet another not-so-great decision. A handful of humans survive in an underwater location in the flooded Earth, and she does find them. Their best plan is to eventually colonize the world ocean with a race of genetically modified humans: a sub-optimal outcome at best.
8. They breach the hull, and most get killed (but not Wilson). The survivors, led by Helen and fellow strong independent woman Holle, end up colonizing Earth III, a cold tidally-locked planet with dense air. Meanwhile the optimists who decided to stay in hostile Earth II do find a way to survive; Baxter later wrote three, separate short stories depicting the future of these colonies.
9. They end up stranded on the universe's least promising piece of real estate for colonization: a chunk of the Moon around the inhabitable Earth. But there's nothing that strong independent women working together without the bother of having men around can't accomplish. You go, girls!
10. Lostetter, to her credit, explores the mythology that such a society generates
11. Not entirely unique, though. In "[Six Wakes](#)" (2017), Mur Lafferty concocts a hibernation ship that, for reasons, is controlled by six convicts who are cloned over and over again (while being injected the memories of their departed versions), while keeping the same roles and positions in the ship.
12. An early wave of despair-driven suicides leads to the first decision to not clone those who killed themselves, since -- the deterministic genetic rationale goes -- that means their genes weren't so ideally suited for the mission to start with. The same happens after, at one point, a rebellion -- led by men, of course -- threatens the mission: the ringleaders are decommissioned and not cloned anymore. Later, the fourth clone of the mission captain kills himself, and his clones are also kept in storage -- so the clone of the second-in-command, the scientist Straifer, becomes captain for a while until, soon after arrival on what turns out to be an actual Dyson Sphere, he goes kind of nuts and dies suddenly. He's also not cloned for a long while. Unsurprisingly, it turns out that cold,

calculating, deterministic scientific types are the most unforgiving people on the universe.

13. Mahler IV kills himself, thus turning his line into undesirable, but is later cloned again to head the Nazi-style squad that keeps an eye over other clones of ex rebels, misfits and suicides while they are worked to death inside of mines in a minor planet captured to extract resources for the mission.

A Libertarian Walks Into a Bear: The Utopian Plot to Liberate an American Town (And Some Bears) by Matthew Hongoltz-Hetling

Out in the storybook land are many leagues of the wild
Wise and uncontainable as a windy-haired child
There are the youth truth movements and they're springing from the ground
Freak flag wavers and you cannot chop them down
Hey, it's the home of the brave, you'll get the threat of the free
Go to the woods
Go to the woods and see
— Dar Williams, "Go to the Woods"

A Libertarian Walks Into a Bear builds on its awesome title¹ in a charming vein: each chapter sports² its very own carefully-chosen literary epigraph about bears, kind of like the humongous chapter of whale epigraphs at the beginning of *Moby-Dick*. The book also holds out the promise of a carefully-researched account of how some libertarian idealism went awry in practice—in this case, a more-radical and smaller-scale predecessor to the [Free State Project](#). Instead of trying to take over the whole state of New Hampshire, the Free Town Project proposed to take over just a single small town there. As you might already guess from the title, that didn't turn out very well³.

For me, the book also delivers on the humor that the punny title suggests; I laughed out loud quite a few times while reading it. Beyond the humor, writing on libertarian experiments gone awry seems significant to me because cases where libertarian ideals have been put into practice are few and far between, and so libertarians can often object that these ideals have yet to be tried. (An earlier contribution in this genre is James Grimmelmann's law review article "[Sealand, Havenco, and the Rule of Law](#)," which examines the downfall of the colo-in-a-microstate from a sort of legal positivist perspective⁴.)

1 Non-native English speakers, please see https://en.wikipedia.org/wiki/Bar_joke.

2 Oops! I mean "bears."

3 I have a libertarian background myself; I wish the Free State Project success, and I don't relish the Free Town Project's failures or feel outraged at its participants' views or anything. In this review, I'll try to describe what I learned from the book, not defend libertarian ideology or accuse the Free Towners of not being Real Libertarians™.

4 You can also place this book in a larger tradition of writing about intentional and utopian communities, where many of what Esperantists might call *samideanoj* choose to move en masse to a single spot and make life the way they think it ought to be. The United States is the undisputed world champion in intentional-community-forming, and there are huge numbers of interesting accounts of the fates of the different communities. Hongoltz-Hetling describes a prior episode in Grafton history in which Unification Church members ("Moonies") also led a mass influx to the town, altering its culture and provoking anxiety among other residents. Other examples today include the several entire Yiddish-speaking towns in New York state founded

I first heard about this book from another review that drew the obvious parallel between “radical libertarians having a hard time coordinating against bear attacks” and “radical libertarians having a hard time coordinating against COVID-19.” And indeed, the question of how to respond to a risk that is *affected*, but not *created*, by individuals’ behavior is a challenging one for libertarian theory. I think it was David Friedman (he can correct me in the comments!) who pointed out that it seems difficult to decide from first principles on the acceptable levels of risk to which we can legitimately expose one another. It’s easy enough to agree that I mustn’t murder you, but [may I fly an airplane over your house](#)? What if it’s an experimental model I built myself? How about failing to finish a prescribed course of antibiotics, and thereby unwittingly giving some bacteria a tiny chance to evolve resistance faster? (They’ll eventually evolve it no matter what, but maybe you just gave them an extra 1,000 shots-on-goal out of 1,000,000,000,000 that they needed, or something?)

Diffuse externalities of many kinds crop up where discrete harmful outcomes are hard to connect to remote, possibly widespread risky individual actions. In the disease case, clearly I can’t deliberately infect you with a contagious disease (that’s widely recognized as a serious form of assault!), but exactly how careful do I have to be once I know I have it myself? How much [effort do I have to take to find out](#) if I have it? How careful should I be to avoid contracting it in the first place because of the subsequent risks to others? All of these questions apply to every contagious disease—even outside of pandemic situations.

These questions mostly lack obvious [Schelling point](#) solutions⁵. Instead, they’re likely to be answered by other means, like gut feelings, or tradition (which doesn’t help much at first for the antibiotics and airplanes), or political processes, or insurance companies’ actuarial assessments (maybe against the backdrop of legal liability rules), or whatever. The answers produced in these ways tend to smack of arbitrariness, which is particularly frustrating to people like me—and to some of the libertarians who moved to Grafton, New Hampshire.

So, Hongoltz-Hetling’s main story begins a decade or two before anyone had ever heard of COVID-19. Lots of radical libertarians, meeting each other online, decide to move to the town of Grafton, described by some as a hospitable environment because of its existing culture (anti-tax, pro-gun, live-and-let-live). When they get there, they’re surprised that the residents are skeptical and worried about them. But why? Shouldn’t they be greeted as liberators, or something?

by Orthodox Jewish denominations and, of course, the Amish and other Anabaptist communities that have tried to live mostly apart from the larger society for centuries.

⁵ A possible counterexample is tort liability for actual injuries, possibly coupled with a requirement to carry insurance against such liability. But this is easier in cases like car crashes than something like transmission of a disease, where direct causality and responsibility may not be apparent at all. It’s certainly not typical for the estates of people who die of transmissible diseases to try to identify and sue the individuals who served as the vectors of their infections.

Well, most of the new residents are conspicuously more radical and explicit about their libertarian views than the existing residents, who are more culturally libertarianish than advocates of a clearly-articulated ideological view. The newcomers want to legalize various things that the old-timers are uneasy with (prostitution, drugs, duels to the death), and want to eliminate every tax possible. The old-timers are largely proud of the relatively small and inexpensive local government they've already achieved and don't specifically think it needs to be shrunk further.

Various things soon go awry. (Hongoltz-Hetling is a local journalist who suggests he stumbled across some of these, in part, while conducting interviews for a different story.) While *A Libertarian Walks Into a Bear* happily chronicles several of the Free Town project's misadventures, the front-and-center one is all about bears and property rights. Basically, the newcomers (and, to a lesser extent, some of the old-timers) have radically divergent intuitions about how to deal with their increasingly noticeable fuzzy forest neighbors. Levy taxes to hire animal-control personnel? Allow town residents to shoot the bears? Get someone from the state wildlife agency to trap and relocate them? Enforce rules about storage and collection of garbage? Have people who are concerned build their own walls or other bear deterrents⁶? Try to adopt them as pets?

The most obvious problem, shown through drawn-out, sometimes hilarious contrasts, is that some of these strategies require more coordination among the town's residents than many of the newcomers are politically comfortable with. While the Free Towners *felt* they were all basically on the same page before they moved to Grafton, their experience of living there together seems to devolve into asking five libertarians how to respond to a problem in a libertarian way and getting six different answers. In what's probably the most extreme case, one animal-loving libertarian resident becomes a sort of Crazy Bear Lady, spending what amounts to thousands of dollars to feed the bears because they're cute and she can tell that they're really hungry. Meanwhile, her neighbors have other ideas, and also, their pets are getting eaten by the bears.

The town also struggles to maintain its fire department. Notwithstanding the existence of libertarian ideas about how to fund firefighting services effectively without taxation, the newcomers don't really succeed in putting any of these proposals into practice, and they also get mad at the libertarian volunteer fire chief—who had helped propose that they all move to Grafton in the first place!—when he tries to enforce state fire codes. In Chekhov's Gun fashion, you have to wonder why the author is mentioning these issues...

Eventually, several tragedies occur, some of the newcomers leave for various reasons, and the Free Town Project pretty much breaks up. Meanwhile, the Free State Project people move to New Hampshire in larger numbers, win a number of elections, and succeed in getting the state to pass a slate of more moderate libertarian reforms⁷. In

⁶ Some people who are building a politically eclectic kind of squatter camp or temporary autonomous zone in the nearby woods actually *do* start to put up some anti-bear walls. They aren't all that effective, but I feel like you just want to cheer for the forest dwellers' efforts. Or, in any case, at least refer to their walls as bearicades?

⁷ The chronology gets a little muddled here. The book made me think the Free State Project migration and electoral successes were all *later* than the main events of the book, but Wikipedia suggests they were mostly

Hongoltz-Hetling's account, these more moderate policy changes don't seem to have any catastrophic effects, and aren't quite as upsetting to other New Hampshire residents, though real controversies do exist and maybe the Free State Project will lead to future political and social divisions. Also, he finds that bear-human conflicts are increasing all across the state, which could be due to increased challenges in coordinating wildlife management policies exacerbated by Free Staters' reluctance to rely on the government to do this coordination.

The author repeatedly expresses his view that funding some kinds of services through taxes can achieve virtuous cycles and produce externalities that make most taxpayers feel they were a good deal. He mostly tries to show this by comparing Grafton with neighboring or nearby New Hampshire cities that have different levels of tax-funded services, pointing out that the cities with more public services got hard-to-quantify but obvious benefits from those services and that most taxpayers were pretty happy with this method of funding them. This isn't a specific theory of how large the State should be or how high tax rates should be, or anything; it's more of an impressionistic observation that radical libertarianism doesn't seem to beat "random small-town New Hampshire fiscal policy" at producing satisfied populations. I took this as a suggestion that people in various places may have levels of taxation that they're culturally satisfied and comfortable with, and be resistant to outsiders wanting to change those levels either upward or downward. (Also, though libertarians have lots of creative ideas for institutions that can fund public goods voluntarily, those institutions themselves don't just magically pop into being without anyone's efforts!)

Relatedly, the book makes interesting and recurring use of the concept of "invasion" (juxtaposing different moments in which human cultures in or around Grafton have been destabilized or disrupted by human or ursine outsiders), which I found thought-provoking. It's a reminder that many ecological and cultural equilibria are possible, and that there are even more possible ways of disturbing them. Chapter 8, while it feels a bit disconnected from the book's main themes, also draws a striking contrast between Grafton and a wealthy New Hampshire town (Hanover, home of Dartmouth College) whose citizens were also distressed by bear incursions in the same timeframe, but managed to use their wealth and political connections to have the bears removed without violating anyone's ethical or aesthetic sensibilities.

In Grafton, public opinion had split between shooting and not shooting the bears. In Hanover, the schism was characteristically different—some people wanted the government to spend a lot of money to modify [a local bear celebrity]'s behavior, while other people wanted the government to spend a lot of money to capture and relocate [the celebrity] and her cubs to someone else's backyard.

It being Hanover, both sides got what they wanted.

contemporaneous. However, Hongoltz-Hetling focuses on the "trigger" declaration in 2016 in which a Free State Project founder publicly called for everyone who was part of the project to act on their pledges to move to New Hampshire. I think the Free State Project had already accomplished a great deal before that and the 2016 event was sort of a declaration of victory.

The contrast is partly about class, partly about ideology, partly about power, and all in all extraordinarily bizarre. (The celebrity mama bear, relocated 1,000 miles away⁸ after Dartmouth elites call in favors, subsequently *walks back* to Hanover, apparently because the food is better or just because she likes people's kindness there. Somebody creates a Twitter account to channel her thoughts on the matter. The government's expenditures and efforts continue to ensure that this story is cute and humorous rather than tragic, finessing the humans' disagreements, but of course "state and local officials can't afford to leverage that kind of effort everywhere," so it's not exactly a prototypical tale of coexistence.)

Back in Grafton, the author also manages to find humor, sometimes fairly dark humor, in all sorts of strange and unhappy events, even including several occasions when people obliquely threaten to murder him⁹. His treatment of the local gun culture, among libertarians and non-libertarians alike, is pretty funny; it's not exactly sympathetic, but not super-judgmental either. There are amusing scenes in which people emphasize how unlike those crazy Free Towners they are, and then talk about their own deep love of and/or willingness to use firearms. If you have a strong view about gun culture, you'll likely be annoyed at the book's equivocal failure to agree with your view.

It would be helpful if the book were clearer on the extent to which the bear chaos can be laid at the paws of the libertarians' actions and political beliefs. The book covers a period of about seventeen years, during which a lot of complicated things involving both bears and libertarianism happened in New Hampshire, both at local and state levels. In some places, the author emphasizes that bear population dynamics and behavior depend on many factors, including weather patterns, human settlement patterns, and actions and decisions by state government (which he notes did not seem to develop adequate capacity to respond to wildlife issues, even before the libertarians showed up en masse). While some Free Towners clearly actively encouraged bears to live in Grafton—like by feeding them—and libertarian culture and politics reduced the capacity for collective or coordinated action on wildlife issues, there are like eight other things that also increased the danger of bear attacks in the same timeframe.

One reason the chronology and causality are so hard to pin down here is that Hongoltz-Hetling is apparently mainly a newspaper journalist who originally mostly did these interviews and research for newspaper-style human interest stories. Newspapers are famously not that great at making detailed arguments about causation or giving you useful context for what you're reading about, compared to reporting on an unusual event or interesting personality. A big limitation of this book, then, is that several chapters feel more or less like extended (often hilarious) newspaper profiles of some individual Graftonite. Good for understanding how that person thinks, or his or her life

8 Other sources suggest it was more like 100 miles, but I can't resist my mental image of the bear singing: "Just to *be* the bear who walked a thousand miles / To fall down at your door..."

9 At one point he starts investigating an incident in which a group of town residents—not necessarily all or mainly libertarian newcomers—reacted to a bear attack on a human by forming a clandestine hunting posse and illegally killing a substantial number of hibernating bears, in clear violation of New Hampshire state hunting regulations. The reaction to his inquiries about this is pretty sinister, as various people imply in different ways that the hibernating-bear poachers might be happy to switch to nosy-journalist poaching.

story; not so great for a synoptic view of the overall impact of libertarianism on bears, or the magnitude of libertarians' role in exacerbating pre-existing problems. Later chapters also often reach back in time, through interviews, to events before those of earlier chapters, or even before the start of the Free Town Project.

I suppose the state-capacity part that the author emphasizes through the wildlife management issues is the flip side of [Bastiat's famous "What is Seen and What is Not Seen"](#) argument, a classic libertarian thought experiment about opportunity cost. Bastiat is totally right that, if the government takes your money to subsidize a gilded bear effigy, people can see the shiny monument in the middle of town, yet not see the small business that you would have opened with that money, or the cool art project of your own that you would have made, or the jobs you would have created through your spending, or whatever. Conversely, if the government takes your money to relocate real bears a thousand miles away from the monument, you *can't* see the absent bears. They're not lolling about wearing "Your Tax Dollars At Work" collars. They're somewhere else!

In the legendarily difficult text-adventure version of *The Hitchhiker's Guide to the Galaxy*, there's an item called "[no tea](#)" which you begin the game with, and can later drop and pick up. (It becomes the basis of one of the game's puzzles.) When you have it, it shows up in your inventory list ("You have: ... no tea"). Much of the humor comes from making explicit one of the things that we *don't* have in our "inventories" most of the time. So, if someone goes to some trouble to discourage bears from hanging out on your doorstep, you might not spend much time pondering the "no bears" item in your inventory, any more than you tend to ponder how you don't have gold bars, or the Mona Lisa, or tigers, or penguins. These may all just feel like basic facts of life. But, depending on where you live, acquiring that coveted "no bears" item may well have taken a lot of unnoticed effort.

I've appreciated the observation that earlier generations were terrified of contagious diseases like smallpox and polio that have been virtually wiped out by vaccination campaigns, while today these diseases are so rare and unfamiliar that many people no longer viscerally fear them or even feel any familiarity with them. That in turn makes it unclear at some intuitive level what it is the vaccines are protecting against: if you don't know *anyone* who has had polio, are you going to feel the same sense of [Jonas Salk as a hero of humanity](#) as someone like my grandmother, who anxiously followed polio outbreak news while worrying about how to protect her children? Are you going to have a clear intuition for why it's important to vaccinate against it? Measles, to take a [tragically contemporary example](#), does feel kind of unreal and unthreatening if you've never encountered it even once in your entire life. Wikipedia [summarizes](#) this observation, made by many medical and public health professionals:

The overwhelming success of certain vaccinations has made certain diseases rare and consequently this has led to incorrect heuristic thinking, in weighing risks against benefits, among people who are vaccine-hesitant. Once such diseases (e.g., Haemophilus influenzae B) decrease in prevalence, people may

no longer appreciate how serious the illness is due to a lack of familiarity with it and become complacent. The lack of personal experience with these diseases reduces the perceived danger and thus reduces the perceived benefit of immunization.

Much the same logic applies to measures protecting against other kinds of threats. [G. K. Chesterton](#) and [Joseph Henrich](#) remind us: that thingy over there might be a weird and broken vestige of history, but it also might *do something we personally can't envision* that responds to *something we can't see*. Nothing makes a problem or danger feel imaginary or hypothetical like handling it successfully!

I recommend this book to libertarians interested in pondering either the theory or practice of libertarian ideas¹⁰. You'll surely wish Hongoltz-Hetling had been more charitable to some of his subjects, or more interested in the substance of their ideas and values. (At one point he casually says Ayn Rand's novels are the prototype for how Free Towners think society should work, even though he's just finished interviewing several of them and seen how their views differ.) You'll probably laugh at some of the wacky hijinks and maybe hold your breath through some moments of terror. (The description of the experience of making eye contact with bears *really* doesn't make you want to go try it out.) If you feel confident in a particular libertarian ethical or political theory, you'll probably have advice ready for the right way that the Free Towners *should* have done things. You may grumble that the book rushes too quickly past the many other factors that contributed to the dangers it describes, most of which were present in Grafton long before modern political libertarianism. But in any case, you'll have lots of things to ponder about how our shared intuitions and understandings bear or don't bear contact with unforeseen circumstances, and how institutions and culture facilitate human life in a world that's not always warm and fuzzy toward us. Or one that can sometimes be a little *too* warm and fuzzy.

¹⁰ If you're, say, a secular urban European social democrat, I fear you may think every single character in the book, including the "moderates" who express skepticism toward the Free Towners, is stark raving mad.

An Evolutionary Theory of Economic Change by Richard R. Nelson and Sidney G. Winter

There's a somewhat straw-man view of economics where we take all the mathematical assumptions at face value. We imagine, for instance, Apple designing the next iPhone. First, they take all of their collected data and build detailed probabilistic models of their customers. Then they consider every possible iPhone design, plug each one into the customer-model, and simulate customers choosing whether to buy the phone or whether to make purchases on it. The company calculates their expected profits and risk, plug those into a model of their longer-term finances, and finally choose the design with maximum long-term expected profit. This is how idealized companies behave in economic models: they calculate a probabilistic model based on all the information available, then choose the option with maximum expected long-run profit. We'll call this the "idealized profit-maximizer" model.

It has been observed, from time to time, that actual companies don't really seem to behave this way.

What actually happens is that Apple takes a bunch of noisy data, builds some crappy models that everybody is correctly skeptical about, then some executives eyeball a few graphs and argue a lot and make a largely-random decision about the final design.

One of the economists' main responses to this observation is a selection argument. In the context of companies maximizing profit, the argument says that companies whose behavior just so happens to maximize profit will end up dominating the industry. They'll have more money to expand faster, invest more, advertise more. It's not that any particular company is a perfectly rational idealized profit-maximizer, it's that the dominant companies act like profit maximizers would, because that's what the market has selected for. It's like fitness-maximization in biology: bees and mushrooms are not perfectly rational fitness maximizers, but they act like fitness-maximizers would, because that's what evolution has selected for.

It's not that Apple figured out which design will maximize profits. It's that Apple *stumbled* on a design which just-so-happened to roughly maximize profits, and as a result they expanded rapidly and that profit-maximizing design took over the industry.

We'll call this picture the "evolution" model, in contrast to the pure "idealized profit-maximizer" view. The key hypothesis of the model is that evolutionary pressure produces (approximate) profit-maximizing behavior, even when companies themselves are not idealized profit-maximizers.

The central question of this review is: does this evolutionary hypothesis actually hold?

What are the hidden assumptions, the approximations, the edge cases? When, and to what extent, will market selection pressure actually result in an industry which behaves like profit-maximizers? How well does the evolution picture hold up when we actually do the math?

Based on Nelson and Winter's 1982 book [An Evolutionary Theory of Economic Change](#), when we do the math the evolution picture holds up remarkably well. An evolutionary model produces (approximate, local) profit-maximizing behavior surprisingly reliably, under fairly weak

assumptions. Nelson and Winter demonstrate this via both proofs (under at-least-sometimes-realistic assumptions) and simulations (under models intended to capture more real-world complexity).

We'll briefly talk about their results - both their main proof and some simulation results - then zoom back out and talk about the broader approach and worldview behind those results.

On the "proofs" side, Nelson and Winter's core argument goes like this: * Businesses change their behavior if-and-only-if profits are low (relative to the rest of the industry) * Profit-maximizing behavior will never have low profits (relative to the rest of the industry) * Therefore, any company which happens to *try* whatever actions are profit maximizing will *continue* to perform those actions indefinitely: they'll never have low profits, so they'll never change. * Conversely, any company with less-than-average profits will try new things until their profits are no longer low. This will push average profits up over time, until the industry converges to profit-maximizing behavior. Some of this sounds unrealistic, but as with any proof we can add more realism with a few simple patches, at the cost of more complexity. The main part which can't really be patched is the idea that a successful company will never abandon their successful strategy. That strains credulity, and relaxing the assumption to something like "businesses *mostly* don't change behavior unless profits are low" completely breaks Nelson and Winter's proof. That said, just because a more realistic assumption wrecks their *proof* doesn't mean that businesses won't still end up as profit-maximizers, it just means more complicated math... which is where the simulations come in.

For people skeptical of the relevance of proofs to complicated real-world systems, the simulations are the real evidence in this book, and they are the bulk of the technical results presented. Nelson and Winter simulate companies with many production techniques, various policies on which technique to use, various research and reinvestment rules, etc, with parameters chosen to roughly match reality. **All of the simulated companies follow randomly-chosen rules, without any explicit profit maximization.** Do the companies which happen to adopt profit-maximizing rules end up dominating the industry? Does the industry overall end up behaving as though companies are profit-maximizers? In chapter 9, Nelson and Winter fit an idealized profit-maximizer model to the simulation results to find out. Their answer: "by and large it seems that [an economist testing the idealized profit-maximizer model] would believe that his theory had performed well".

That's a rather spectacular understatement - they fit an idealized profit-maximizing-firm model to the evolutionary simulation data and find an R^2 value of 0.99. In other words, companies subject to selection pressure approximate profit maximization ridiculously well; evolution does indeed induce near-ideal profit-maximizing behavior. The evolutionary hypothesis holds extremely well.

But is this result really that impressive? Presumably Nelson and Winter set out to build a foundation for orthodox profit-maximizer-style economic models; of course they would choose parameters for which the evolution-induces-maximization model works well.

Except that Nelson and Winter weren't trying to build a foundation for profit-maximizer models. What makes the results really believable is that they were quite explicitly trying to *attack* those models.

Habit and Routine, Not Calculation and Optimization

Nelson and Winter have a vision of how economic theory ought to look. It's a vision in a similar vein to James Scott ([Seeing Like a State, Against The Grain](#)), Michael Polanyi ([The Great Transformation](#)), or Joseph Schumpeter. Knowledge is not explicit, it's tacit. Local and contextual, not global and general. Habits, not calculation. Sometimes the steel mill's process only works because of an impurity in their ore which neither the mill nor the supplier knows about. One worker always leaves tool A in bin B, another worker always looks for tool A in bin B, and nobody ever thought about it because it just works. Maybe someday a new worker will come along and leave tool A in bin C, and other workers will be confused for a bit until they settle into the new normal. People mostly keep doing what works, and if anything goes wrong they'll look around for something else which works and then do that instead.

Businesses, in this vision, are the same thing at a larger scale. They hire people, adjust processes, introduce new products or pull old products, generally try stuff out and see what works. They don't know that the sudden uptick in sales of an old sunblock product is due to the opening of a new waterpark, they just know that sales are suddenly up so they better make more of it. If the waterpark shuts down for some reason, the company won't know that they need to cut back production until a backlog of sunblock builds up in the warehouse. The business doesn't really know what aspects of the environment are crucial to their profitability, which employees are performing crucial tasks that nobody else thinks about, which accidental features of their products are crucial to downstream consumers.

In a static world, it wouldn't matter, at least for predicting the behavior of businesses. The results from the previous section tell us that businesses will randomly try things until they hit on (approximately) optimal choices, and then mostly keep doing that. They end up acting like profit-maximizers, regardless of their actual decision processes.

But this breaks down when things change quickly.

Take a worker from one company, hire them to do nominally the same job at a different company, and they'll spend a month or two largely confused and ineffective while they figure out where things are, which people they need to talk to when, and all those other little things that don't directly translate from one company to another. Take a moderately successful thirty-year-old family restaurant and throw it into the environment of COVID and delivery, and the restaurant will thrash around desperately trying different menus, prices, hours, delivery and payment options until they either find a combination which works or shut down. There is no methodical enumeration of all available options, no Bayesian update. In retrospect, the restaurateurs may realize that they had all the information needed to identify the best option - but hindsight is 20/20.

In other words, when things change quickly, the evolutionary viewpoint says that businesses' behavior should be drastically suboptimal for some time. Not indefinitely; over time those businesses will figure out what works (or go out of business and be replaced by others). But we shouldn't expect profit-maximizing behavior on short timescales in response to major changes.

Thus the need for an evolutionary theory of economic *change*, specifically. Idealized profit-maximization models, in which we pretend that companies make the best-possible decision every time, might be able to predict a long-run equilibrium. But they can't properly predict the *path* to that equilibrium, because in the real world that path involves companies making less-than-ideal decisions and learning from their mistakes. And in a sufficiently rapidly-changing world, we may never get close to the equilibrium anyway; the path is all there is.

Nelson and Winter expound on such shortcomings of “orthodox” (i.e. idealized profit-maximizer) economic models for about half of the book.

Economics Has Come A Long Way

Unfortunately, while they *talk* at length about the inability of idealized profit-maximizer models to handle high uncertainty, fast change, tacit contextual knowledge, and the like, Nelson and Winter’s math doesn’t really back it up.

This is probably more clear in hindsight than it was in 1982. When Nelson and Winter talk about things which are “fundamentally incompatible” with idealized profit-maximizer models, a fairly large chunk of the commentary is about nonequilibrium dynamics (i.e. fast change) and high uncertainty. The tools to handle those (mainly dynamic programming, Bayesian probability, and game theory) were around in the ‘80’s, but they hadn’t been integrated into the “core” economic models to nearly the extent that they are today. For instance, the more recent models in Ljungqvist and Sargent’s [Recursive Macroeconomic Theory](#) are very clearly in the same idealized profit-maximizer family which Nelson and Winter criticize, yet they directly handle dynamics and uncertainty. The core models of economics just aren’t dependent on static environments and perfect information to the extent that they used to be; those shortcomings turned out to be largely independent of the idealized profit-maximization assumption.

Conversely, the evolutionary models which Nelson and Winter themselves present mostly seem to provide a solid foundation for the idealized profit-maximization viewpoint: despite involving plenty of dynamics and uncertainty, their results mostly look like the simulated companies are quite close to idealized profit-maximizing behavior most of the time. Much as they want to overthrow the idealized profit-maximizer models, they don’t really show evolutionary models behaving dramatically differently from how idealized profit-maximizer models would behave. They expend a lot of ink arguing about how evolutionary models are superior to idealized profit-maximizer models and should largely replace them, but their actual results don’t really back it up.

That said, to a large extent I think this is a failure of Nelson and Winter rather than a failure of the evolutionary approach. Just as orthodox economic theory has come a long way over forty years, evolutionary economics has *also* come a long way - a two-minute perusal of [this 2004 paper](#), for instance, reveals a much more substantive mathematical foundation than Nelson and Winter used, even in a paper which isn’t really doing much sophisticated modelling. That mathematical foundation matters if we want to do things like, say, back-of-the-envelope estimates to figure out whether and when the predictions of an evolutionary model will diverge from the corresponding idealized profit-maximizer model.

I very much doubt that Nelson and Winter’s commentary was entirely wrong; we *should* expect the predictions of idealized profit-maximizer models to diverge from reality under conditions of rapid change. But without a sufficiently solid mathematical foundation, we have no idea *how* fast things need to change, or which things, in order to see novel testable predictions from evolutionary theory. What precisely are the conditions under which the idealized profit-maximization approximation breaks down? That’s the sort of foundation I was hoping to see in Nelson and Winter, and which I still hope to see - but I’m not sure where to look.

Antifragile: Things That Gain from Disorder by Nassim Nicholas Taleb

I. Prologue

[*Antifragile: Things That Gain from Disorder*](#), by Nassim Nicholas Taleb, seems to have anticipated my project of reviewing it, and took pains to make it difficult for me. Writing on the book's structure:

Someone in the business of "summarizing" books would have to write four or five separate descriptions. But I saw that they were not stand-alone essays at all; each deals with the applications of a central idea, going either deeper or into different territories: evolution, politics, business innovation, scientific discovery, economics, ethics, epistemology, and general philosophy.

Books to me are not expanded journal articles but reading experiences; and the academics who tend to read in order to cite in their writing—rather than read for enjoyment, curiosity, or simply because they like to read—tend to be frustrated when they can't rapidly scan the text and summarize it in one sentence that connects it to some existing discourse in which they have been involved. Further, the textbook is the polar opposite of the textbook—mixing autobiographical musings and parables with more philosophical and scientific investigations. I write about probability with my entire soul and my entire experiences in the risk-taking business; I write with my scars, hence my thought is inseparable from autobiography. The personal essay form is ideal for the topic of incertitude.

A noble aim, even if it makes my job a bit tougher. To balance the scales, I too shall intrude on the text and offer autobiographical details of my own, when relevant. Though I will not, in the French style, beg forgiveness from you, dear reader, for such digressions. I impose them of my own accord.

Antifragile is certainly no textbook. It intermittently takes on the form of a morality tale, with clear heroes and villains. Taleb's heroes include grandmothers, powerlifters, mafia dons, William Gladstone, the Stoics, Muslim poets, Christian saints, Ralph Nader, flaneurs, artisans, and Mother Nature. His villains include suckers, naive rationalists,

turkeys, ‘interventionistas’, bankers, Thomas Friedman, career academics, consultants, Myron Scholes, and Ray Kurzweil.

If you were worried that going mainstream might have sanded away his rough edges, rest assured that his eccentricities are back in full force. Taleb’s use of fictional characters has by no means gotten more subtle since *The Black Swan*. The work is sprinkled with the conversations of ruminating-bookreader Nero Tulip and boisterous-steak-eater Fat Tony. At one point, Taleb places Fat Tony in the *Euthyphro* dialogue with Socrates, and has his character trounce the old philosopher by reframing the questions.

And the key element of every Taleb book is present right from the start: everyone but him is wrong.

II. The Triad

The University of Chicago’s [writing program](#) urges its students (faculty and PhDs, not undergrads) to focus on creating value in their work before anything else. How do you create value? If nothing else, say someone is wrong and try to prove it.

Taleb would have passed such a course with flying colors. He accumulates a lengthy list of people who were wrong, whether of the “sucker,” “sissy,” “naive rationalist,” or “interventionista” flavor, which grows to include nearly all economists, great swaths of the business, technological and medical professions, and philosophers all the way back to Aristotle and Socrates.

He starts by criticizing dictionaries and thesauruses—all of them, in every major language—for making the same mistake. They all list the opposite of “fragility” as “robustness,” “toughness,” or somesuch. Taleb is unimpressed.

Almost all people answer that the opposite of fragile is “robust,” “resilient,” “solid” or something of the sort. But the resilient, robust (and company) are items that neither break nor improve, so you would not need to write anything on them—have you ever seen a package with “robust” in thick green letters stamped on it? Logically, the exact opposite of a “fragile” parcel would be a package on which one has written “please mishandle” or “please handle carelessly.”

The objects with this property—benefiting from shocks and randomness— are antifragile. They include all living things (at some level), ecologies, complex systems, information, and some people.

Taleb illustrates a triad, with fragility on the left, robustness in the center, and antifragility on the right, and ties each one to a mythological figure (Damocles-Phoenix-Hydra). Fragility and antifragility are each other's opposites, twins with the signs flipped. Robustness is a null state between them. In models of fragility, these may be explicitly used as parameters from -1 to 1.

Taleb previously observed the effects of nonlinearities in *The Black Swan*, such as with the teacup, which suffers far more from a single drop from one foot than from twelve drops from one inch. But only now has the wider insight become clear. No matter how small the shocks, any object, no matter how robust, will eventually break if exposed to enough of them. It's purely a matter of time, and the universe has nothing but time.

How is it then that anything has survived to the modern day, after billions of years of shocks, some very large? Surely, that which has survived must be more than robust.

This is the key distinction between the organic and inorganic; in Taleb's lingo, the cat and the washing machine. The inorganic will necessarily break, each shock only brings it closer to that end. But the cat can survive very strong shocks and improve. Not only does it heal (say from accidentally crawling inside the washing machine) but there's a good chance it will be *stronger* for the experience, after a suitable recovery period, and provided the shock was below the lethal or crippling threshold..

In this light, shocks and stressors are necessary, not just helpful, since they represent information about the environment. This is true for cats as well as national economies. Naive attempts to cut shocks and stressors out of life creates long-term fragility, as an astronaut is weaker after going without gravity. And rest assured, everything comes falling back to earth in due time.

III. Barbell Everything

Taleb spends approximately half the book going on about asymmetries—especially convexity effects—and their applications to every field imaginable. He's even managed to evolve one of the concepts previously debuted in *The Black Swan*: the barbell.

The barbell, or the less-Brooklyn “bimodal strategy,” places weight on the extremes. In vulgar finance, take extremely safe *and* extremely risky bets, but nothing moderate. But now, Taleb has expanded it to virtually every domain, with a little help from Lucius Annaeus Seneca.

He analogizes the Stoics’ teachings of denigrating possessions to cutting off downside from a bet. The Stoic is at the mercy of whimsical Fortuna, and does not know if he shall be wealthy or poor, successful or a failure. So he mentally writes off anything that he could possibly lose: every material possession, his net worth, even other people. No longer dreading the loss of those things, he is free to enjoy them, and if Fortuna sees fit to take them away, the Stoic feels as if he has lost nothing. He has domesticated Fortuna.

Taleb raises this to an aesthetic, and a full way of life. He writes accessible personal essays and highly technical works, never hybrids. He powerlifts, trying to break his previous record, and takes it easy the rest of the time. He doesn’t smoke, eat processed foods, ride motorcycles or mess with the mob, but he insults entire professions in books and takes nail-biting bets with his portfolio.

In conversation partners, seek out brilliant scientists and taxi drivers, never middle managers. Marry the accountant and have an affair with the rock star. Be an ascetic, then go to the craziest party in town once a week.

This is the domestication of uncertainty, minimizing how much it can hurt you, then opening yourself up to its rewards. When you’re barbelled, you benefit from volatility. When you’re not, you benefit from monotony. And monotony never lasts.

IV. Killing Volatility

Taleb complains bitterly about a particular modern tendency: the privileging of stability and predictability over volatility. The desire to eliminate major trauma leads to a life filled with high-frequency, low-intensity stressors. This drives him up the wall.

Our antifragilities have conditions. The frequency of stressors matters a bit. Humans tend to do better with acute than with chronic stressors, particularly when the former are followed by ample time for recovery, which allows the stressors to do their jobs as messengers. For instance, having an intense emotional shock from seeing a snake coming out of my keyboard or a vampire entering my room, followed by a period of soothing

safety (with chamomile tea and baroque music) long enough for me to regain my emotions, would be beneficial for my health, provided of course that I manage to overcome the snake or vampire after an arduous, hopefully heroic fight and have a picture taken next to the dead predator. Such a stressor would be certainly better than the mild but continuous stress of a boss, mortgage, tax problems, guilt over procrastinating with one's tax return, exam pressures, chores, emails to answer, forms to complete, daily commutes—things that make you feel trapped in life.

There are two ways to destroy an antifragile object. One, hit it with a shock too big for it to recover from. Meteors usually work. Two, prevent it from recovering. It's how Heracles neutralized the antifragile Hydra, and how our ancestors chased down faster prey.

Unfortunately, modernity tends towards exactly this model of stressors. It abhors volatility, and prefers a steady and regular regimen of everything from food to exercise, learning to income. There's no room to recover, and no room to suffer. As a counterpoint, Taleb advocates adding volatility back into your life: fasting and powerlifting are his favorite examples. The signals carry more weight when not drowned out by monotony. Life is more enjoyable with peaks and valleys.

It's also more stable in the long term. Volatile stressors, so long as they remain non-lethal, expose the weaknesses in systems. Meanwhile, attempts to quash volatility, such as in the economy, only hide the signal and compress volatility. The system grows vulnerable, like a forest deprived of fire. When the shock finally comes—and it always comes—it will be worse than otherwise expected, and the system will be poorly prepared. In trying to iron out the boom-bust cycle, one only succeeds in creating a crash worse than any before seen.

Here we come to one of Taleb's recurring villains, the fragilista. The empty suit who, thinking he can improve the system, only moves fragility around, priming the system for destruction. In politics as in economics as in demonology, tampering with forces beyond one's ken has predictable results. They treat the economy as a washing machine, and the state as a sailing ship, as machines which must be maintained by a qualified repairman, instead of primal forces.

Social scientists use the term "equilibrium" to describe balance between opposing forces, say, supply and demand, so small disturbances or deviations in one direction, like those of a pendulum, would be countered by an adjustment in the opposite direction that would bring things back to stability. In short, this is thought to be the goal for an economy.

Looking deeper into what these social scientists want us to get into, such a goal can be death. For the complexity theorist Stuart Kaufman uses the idea of equilibrium to separate the two different worlds of Table 2. *For the nonorganic, noncomplex, say, and object on the table, equilibrium (as traditionally defined) happens in a state of inertia. So for something organic, equilibrium (in that sense) only happens with death.*

Life, especially the most interesting parts, occur in the swing of the pendulum, in the vortex, in the variance, not the mean. Stability is death.

If I could predict what my day would exactly look like, I would feel a little bit dead.

V. Against the Nerds

Of all the people to wind up devaluing academia, you wouldn't have expected Taleb. His father was an anthropologist, and his mother an oncologist. In a resounding win for nominative determinism, his father's name was Nagib Taleb, which roughly translates as 'Student Intelligent' (or so Taleb claims), and he went on to be a national valedictorian in Lebanon. And his mother's name was Minerva, all the more.

Nevertheless, Taleb did not attend the rigorous Lebanese Jesuit schools, and grew to understand the negatives of being an "Intelligent Student," an intellectual deformity that selects for focusing on boring material. He went far in the opposite direction, claiming to be an autodidact in spite of his degrees. He relates his childhood habit of reading any book that wasn't on the assigned list, of doing only enough work to not flunk and spend the rest of his time learning his own way.

Even so, he graduated from The University of Paris and Wharton Business School, the consummate foreign-educated upper-middle-class individual. How did he become so enamored with Brooklyn taxi drivers and mafia dons, and so anti-nerd?

My intellectual world was shattered as if everything I had studied was not just useless but an organized scam—as follows ... I had to cohabit with foreign exchange traders—people who were not involved in technical instruments as I was; their job simply consisted of buying and selling currencies ... Coming to this from a highly polished Ivy League environment, I was in for a bit of a shock. You would think that the people

who specialized in foreign exchange understood economics, geopolitics, mathematics, the future price of currencies, differentials between prices in countries. Or that they read assiduously the economics reports published in glossy papers by various institutes. You might also imagine cosmopolitan fellows who wear ascots at the opera on Saturday night, make wine sommeliers nervous, and take tango lessons on Wednesday afternoons. Or spoke intelligible English. None of that.

My first day on the job was an astounding discovery of the real world. The population in foreign exchange was at the time mostly composed of New Jersey/Brooklyn Italian fellows. Those were street, very street people who had started in the back office of banks doing wire transfers, and when the market expanded, even exploded, with the growth of commerce and the free-floating of currencies, they developed into traders and became prominent in the business. And prosperous.

My first conversation with an expert was with a fellow called B. Something-that-ends-with-a-vowel dressed in a handmade Brioni suit. I was told he was the biggest Swiss franc trader in the world, a legend in his day—he had predicted the big dollar collapse in the 1980s and controlled huge positions. But a short conversation with him revealed that he could not place Switzerland on the map. Foolish as I was, I thought he was Swiss Italian, yet he did not know there were Italian-speaking people in Switzerland. He had never been there. **When I saw that he was not the exception, I started freaking out watching all these years of education evaporating in front of my eyes.** That very same day I stopped reading economic reports. I felt nauseous for a while during this enterprise of “deintellectualization”—in fact, I may not have recovered yet.

That ... ought to do it.

Taleb derives from this that the price and the reality (or even worse, the economic theory) are not the same thing. You can predict war and still not realize that it necessarily means a rise in oil. You can know everything about Switzerland, and have no advantage (or be at a disadvantage owing to overconfidence and spurious details) when it comes to trading francs.

VI. Against the Predictors

Taleb derives a great error in thought from Aquinas, quoting Averroes, “An agent does not move except out of intention for an end.”

So let us call here the teleological fallacy the illusion that you know exactly where you are going, and that you knew exactly where you were going in the past, and that others have succeeded in the past by knowing where they were going.

Knowing where you are going seems a pretty valuable skill. If you're planning a trip you'd like to have an itinerary. If you're running a business, you'd like to have a business plan.

Yet Taleb rejects exactly this; detailed planning beyond general intention is actively harmful. He contrasts the tourist, who is the prisoner of a plan, to the rational flaneur, who changes their plans at each stop (or at least has the option of doing so). He modifies his targets as he gains more information, so the experience is continually refined!

The same goes for the dreaded business plan. Taleb cites a laundry list of prominent companies which began in domains totally disconnected to what finally made them rich; consider how Nokia started as a paper mill, and Raytheon started out making refrigerators.

The key is optionality, and the preservation of optionality. Locking yourself into a single plan prevents you from revising and changing tactics as you gain more information and the landscape changes. It's fragile.

Taleb makes an even bigger claim. Optionality is preferable to *intelligence itself*, or can serve as a substitute for intelligence.

If you “have optionality”, you don't have much need for what is commonly called intelligence, knowledge, insight, skills, and these complicated things that take place in our brain cells. For you don't have to be right that often. All you need is the wisdom to *not do* unintelligent things to hurt yourself (some acts of omission) and recognize favorable outcomes when they occur.

This is Taleb's ‘philosopher's stone’ aka convexity bias. It's also the driving force behind evolution. Every mutation and recombination represents extra options for nature to choose from in selection. All that remains is to have a good filtering mechanism, and that comes for free.

All very convenient for Taleb, whose central thesis is “how to live in a world I don’t understand.” He might like the post-rationalists. Fragile systems require a great deal of intelligence to keep running. You need to accurately predict the future, *and* prepare for it. Taleb cites oil speculators in the leadup to the US invasion of Kuwait, who spent immense resources predicting if war would break out. In fact, they predicted war accurately, but not its consequence. Oil prices dropped instead of rising, and many sophisticated predictors lost their shirt. Meanwhile, Fat Tony observed all this from the sidelines, decided that they were all suckers, and made fat stacks without knowing the first thing about Kuwait, just by betting against the fragile predictors.

If you benefit from convexity bias—if you’re antifragile—you only need enough intelligence to avoid catastrophic risk and to pick up your shekels. The rest takes care of itself.

VII. Via Negativa

Seven sections into the review, and the core still eludes me. Well, instead of talking about the book, let’s talk about what the book isn’t. In Taleb’s own words, it’s non-sissy, non-predictive, non-turkey, non-interventionist. I suppose one could read it in a nonlinear fashion.

This method of talking about what a thing isn’t, *via negativa*, is a tool Taleb picks up and uses on everything. Rather than adding, he advocates subtraction wherever possible. Subtracting from your diet, from your doctor’s visits, from your reading list, from your portfolio, from knowledge. Remove errors first, and watch problems evaporate.

While in the emergency room for a swollen nose, Taleb grew suspicious of treating swelling with ice, and asked the doctor if there were any studies showing the effectiveness of ice therapy. The doctor replied:

You have a nose the size of Cleveland and you are now interested in ... numbers?

To Taleb, this is an example of blind intervention, as his swollen nose clearly wasn’t an instance of danger from the swelling. The inflammation ought not be treated, as it is a necessary step in the body’s healing process.

But this is merely unwise and inconvenient, not dangerous. The unnecessary ice patch is the foundation upon which the edifice of iatrogenics (harm by the healer) is built! Taleb goes as far as to say that, if you want to hasten someone's death, pay for them to have a personal doctor. They don't have to be murderous, or even bad, simply having access to medical treatment encourages using it, and every additional procedure and pill removes one from Mother Nature's time-tested methods.

Taleb would have you subtract from your life everything that hasn't been around for at least a thousand years, or which doesn't remove the effects of other technology. Wine and coffee are fine, but sodas aren't. Digital technology is as well, because it removes inconveniences from the earlier modern and industrial eras. I can't for the life of me see how Twitter meets this criteria, but I'm sure he has an elegant explanation elsewhere.

VIII. Against the Immortals

So, is there a great difference between Taleb and the rationalists?

On the surface, maybe not. He critiques the Nerd as someone "who thinks exceedingly inside the box" and the Naive Rationalist as "he who thinks the causes for things are knowable, and known to him." Both are failure modes explicitly called out by the modern rationalist community.

Taleb, the pit trader-turned fund manager, rejects "verbalistic" pursuits, where being right is a matter of convincing others, and prefers instead to use correct knowledge to make money. He would probably approve of bet prediction sites and the rationalist virtue of making bets.

On the other hand... Taleb reminds me of Dumbledore. [Eliezer Yudkowsky's Dumbledore](#).

And, of course, we have this modern illusion that we should live as long as we can. As if we were each the end product. This idea of the "me" as a unit can be traced to the Enlightenment. And with it, fragility.

I was just reading in John Gray's wonderful *The Immortalization Commission* about attempts to use science, in a postreligious world, to achieve immortality. I felt some deep disgust—as would any ancient—at the efforts of the "singularity" thinkers (such as Ray Kurzweil) who believe in humans' potential to live forever. Note that if I had to find the anti-me,

the person with diametrically opposite ideas and lifestyle on the planet, it would be that Ray Kurzweil fellow. It is not just neomania. While I propose removing offensive elements from people's diets (and lives), he works by adding, popping close to two hundred pills daily. Beyond that these attempts at immortality leave me with a deep moral revulsion.

It is the same kind of deep internal disgust that takes hold of me when I see a rich eighty-two-year-old man surrounded with "babes," twentysomething mistresses (often Russian or Ukrainian). I am not here to live forever, as a sick animal. Recall that the antifragility of a system comes from the morality of its components—and I am part of that larger population called humans. I am here to die a heroic death for the sake of the collective, to produce offspring (and prepare them for life and provide for them), or eventually, books—my information, that is, my genes, the antifragile in me, should be the ones seeking immortality, not me.

Then say goodbye, have a nice funeral in St. Sergius (Mar Sarkis) in Amioun, and, as the French say, *place aux autres*—make room for others.

It's fair to say Taleb isn't onboard with the transhumanist mission.

I intrude here. How is it that two of my greatest influences, Taleb and Yudkowsky, who have both offered me insights which have transformed my life, could be at such odds, not on any trivial matter, but perhaps the most important?

I foresee the simplest answer: Taleb is completing a Deeply Wise™ cache native to his Greek Orthodox background. Yet, I'm not comfortable writing him off like this. His belief flows from his other positions and arguments smoothly. He expresses elsewhere that nature's abhorrence of the individual, the evolutionary necessity to sacrifice the organism or the long-term survival of the species, horrifies him. Yet he admires the primacy of nature too much to offer it any defiance.

This is where Taleb and I must part ways.

IX. Autobiography

What more can I say? I reject the vulgar practice of assigning scores or stars to reviews. If my above description of the book's contents and arguments don't convince you to read it, I can hardly imagine what will. So I will indulge, and write about myself.

I first read *Antifragile* in the summer before my junior year of high school (actually, I listened to the audiobook). Growing up, I was the consummate nerd, in Taleb's vocabulary. Extraordinarily risk-averse, anti-social, eager-to-please to a destructive extent, and widely dismissive and resentful of others.

Antifragile wasn't the book that burst me out of that bubble; it was *Twelve Rules for Life*, whose sequel is actually coming out soon. But *Antifragile* was the next major source of insights, confirming much of *Twelve Rules* for me while encouraging me to move on from the Golden Lobster Idols. In particular, *Antifragile* is the book that got me into mathematics.

I didn't grow up as a 'math person.' At a young age I could do mental math quite well, but the insistence on *writing out* what I could do in my head just fine antagonized me to such an extent that I refused to value the subject. I conceived myself as an essayist, or maybe a writer.

Then came a book which seemed to explain so much, and backed up its arguments with statistics. Not much statistics mind you, but simple, accessible notions. I dug further into Taleb's writing, and began to idolize him. For the first time, the beauty of mathematics and what it could do were apparent. That year, I took Pre-Calculus, and my instructor (an acerbic, humorous teacher whom I loved and my best friend hated) stoked that flame even hotter.

It wasn't long after that I first read the Scott's *Meditations on Moloch* (I wrote a poetry chapbook riffing off it my senior year) and the *Sequences*. Without Taleb and that professor to bridge the gap and give me an appreciation for quantitative fields, I might have bounced off rationality entirely.

What more can a book do for me? When writing my college admissions letter, I reached back to Taleb and syncretized his notions of antifragility with the school's mascot, the Phoenix (a little white lie—the proper representation would be Hydra, which isn't as sexy—but they can't un-admit me now!). For all I know, that tipped the scales towards admission, and I now study statistics there.

More? I lent the book to a girl my freshman year of college. That didn't go anywhere. I suppose expecting a book to get me hookups is a bit too much to ask. Still, very useful.

Scott already mentioned in his review of *The Black Swan* that Taleb was trying to revive certain archaic ideas about the intellectual with Greek and Roman names. Here's another for that list: *megalopsychon*, the great-souled man, who takes risks and remains dignified in the face of all things, before whom the words of non-risk-taking suckers are as the sounds of animals!

The arrogance! I adore it! It is a concept, a stance for which I have cried out without knowing it. The hero who is high! The unbridled ego of Taleb! The confidence to stare down whole professions!

Taleb! Taleb! Fragile economies! Wicked doctors! Steak dinners! Partial veganism! Suckers to be fleeced! Saints to be praised! Consultants weeping in the audience!

But I get ahead of myself—ecstatic beat poetry is not appropriate for a book review.

I collect the voices of my favorite authors and let them battle it out in my head. He and Peterson have such wonderful discussions in there! I put them on as masks too. The Taleb mask has served me well. When I need confidence, when I need to face down a narcissist, when I need to defend my worldview from the enemy, I reach for him.

If you too want to have the voice of a Lebanese volatility specialist in your ear, judging every aspect of your life, read [*Antifragile: Things That Gain from Disorder*](#).

Anna Karenina by Leo Tolstoy

I.

Isn't this just some long-winded Russian book about adultery and divorce? Why are there legends about Tolstoy and about Anna Karenina?

One of the first [stories](#) I heard about Anna Karenina was from Teddy Roosevelt. He was rudely interrupted from hunting mountain lions in the Dakotas when some common cattle-killing horse-thieves ran off with his boat in the night. Roosevelt built a raft, chased them down through the still-icy river and captured them. Rather than hanging them where he caught them as was the custom, he decided to march them 300 miles to the sheriff of the nearest town while he read Anna Karenina.

Then there was Tolstoy's personal correspondence with Gandhi, who called Tolstoy "the greatest apostle of non-violence the present age has produced."

Then there was social activist Jane Addams's journey to Russia to seek Tolstoy's spiritual counsel, which began next to the peasants in the hayfields, with Tolstoy laying down his scythe and "pulling out one of the big sleeves of her fashionable traveling gown, saying it contained enough fabric to make a smock for a little girl." This biographical blurb about Jane Addams continues with: "Tolerating this and other discrediting jibes during the encounter, Addams returned to her demanding schedule at Hull-House, concluding that strict adherence to Tolstoy's philosophy would be counterproductive and 'utterly preposterous.'"

Other contemporaries like William James [wrote](#) about Tolstoy's philosophy, saying "he makes his love of the peasant so exclusive, and hardens his heart toward the educated."

Those last anecdotes are the most revealing about what's in a Tolstoy novel.

Why is there so much myth and divisiveness surrounding the man who wrote about adultery in Anna Karenina and love triangles in War and Peace? After seeing the scope of Anna Karenina, I understand. Teddy Roosevelt wasn't really reading a soap opera about a sexually-awakened housewife as he marched thieves across the Badlands. He was reading a sprawling novel about Russian upper class life from the one other aristocrat who would famously prefer to tromp through the country with whichever men he encountered rather than attend to high society.

Anna Karenina is a 346,000-word effort of empathy from one helplessly offputting man. Okay, maybe it's more like a 100,000-word effort, interrupted by 246,000 words of personal politics and contempt, but that's altogether a more thorough effort to respect more varieties of people than I've made in my life. Tolstoy was a social conservative trying to accept why a woman might reasonably leave her family for lust. To mirror Anna's struggles, he added an co-protagonist modelled after himself in his final bachelor years, who intermittently reads like a self-mocking apology to everybody who has had to interact with Tolstoy and particularly to those who had to love him.

Tolstoy does showcase his divisiveness throughout the novel, the divisiveness that drew Jane Addams first towards then away from him. I don't think he's actually more offensive than any other partisan writing to their ingroup about an outgroup. The difference is that for Tolstoy, the outgroup is virtually everybody who has the time or pretense to read books rather than eternally struggle with the land for a meagre existence. So it's a testament to Tolstoy's intermittent empathy that anybody likes the novel.

II.

First, about that adultery. How do these 19th century moralists write the sex? The author of *Lolita* described Russian literature as "altogether the most chaste of novels." In the Puritan novel, *The Scarlet Letter*, the sex happens before page 1. Tolstoy can't do that with *Anna Karenina*, because he wants to show her descent into sin. Here's her passionate tryst in full graphic detail:

That which for Anna had been an impossible, horrible, but all the more enchanting dream of happiness--this desire had been satisfied.

Anna's tryst is the reversible beginning of her fall from grace. What ruins her in Russian high society is she stops hiding her affair; she becomes not only gossiped about but also seen in closer relations with her lover than her husband. In the second half of the book, we hear very infrequently from Anna herself, but more of how other people deal with her and think of her. She's miserable, and her defiant existence makes everyone uncomfortable, including I confess myself. It was a sorry relief to me when one 100 page section offered more point-of-view sentences to Levin's hunting dog than to *Anna Karenina*.

You can read a lot about Anna from other sources, so I'm going to focus on some other parts of the book.

III.

There are no villains. Because the book has an average of 100 pages per POV character, I didn't keep the same opinions of them throughout the story. Sometimes they are easy to hate (just like the author), but other times they are endearing, and then shortly afterwards bear an uncomfortable resemblance to myself.

In a later Russian book (*The Gulag Archipelago*), the phrase "the line between good and evil cuts through every human heart" The Russians are great at stopping and making you look uncomfortably into yourself, but that only happens after you read them for a while and jump past the offensive attitudes.

Tolstoy makes me question what other people think of me; Dostoyevsky makes me question what God thinks of me. Not exactly God, as an atheist, but that's the most succinct way to put it, just whenever the reasons I tell myself for doing something sound more noble than why I might really be doing it. And "what other people think of me" is not necessarily the best metric of whether I'm fooling myself.

Anna's brother is a schmooze, but he has his time to shine when he hosts a dinner party. His ambition, other than chasing skirts, is to curate social situations where everybody delights in each other's company and any uncomfortable hatreds are swept away. This sets him up to be what another reviewer described as the Master of Ceremonies of the novel. His betrayed wife says of him "I love him, but I do not respect him," and I feel the same way. In one scene, he is making house calls in preparation for a dinner party in which he wants to, in addition to the usual things, play matchmaker between his sister-in-law and his childhood friend, and persuade his brother-in-law to stay with his sister. In between dinner party planning (and remember there are servants to take care of the less-important parts like cooking and cleaning), he stops for a three-hour lunch with his new uptight boss, whom he loosens up for the benefit of himself and all his coworkers, while making a point to wear the Russian aristocracy equivalent of a t-shirt.

Anna's husband...might not get much sympathy with we SSC crowd. He's more concerned with politics than anything else, but then becomes a devout Orthodox Christian just in time to be maximally cruel to everybody around him. Christianity was a bigger deal in the high-society circles back then. But he becomes more sympathetic when you substitute "Christian morals" with [any other morality system], where when you decide how to treat somebody you are balancing whatever are their desires with any other moral guidance for what you think is best for them.

There were two moments where I noticed Karenin's humanity and sympathy: his discomfort with tears, and his response to his son's pride that he won some political award.

You can look at Anna's husband with contempt the whole way through the novel, as some Christian oaf, but I also began to sympathize with him and also his son and wife who can't stand him. He looks at dashing young cavalry officers and political awards and thinks they're dumb, occasionally wondering what his wife and son see in them. You also see how he behaves at the horse races, immediately finding some other intellectuals to argue with rather than attending to the race itself, saying things like "the love of such spectacles is the surest sign of low development in the spectators." I have this vision of meeting modern-day Karenin at a Super Bowl party, spending the whole time nasally nattering about how football is stupid and a Bad Influence on society, with the degenerative brain disease, the cheerleaders, the light beer, the vitriol between the two tribes of fans, and did you know how many anti-malarial bed nets you can buy for the cost of one Super Bowl commercial, while everybody else says "shut up, we know it's dumb, but we like it and you're stifling us!" Anna's complaint of her husband was that living with him was stifling, and she used the word "magnaminious" as a slur in describing him. The original 1870s Russia Karenin's moral awakening was to Orthodox Catholicism, but the 21st century American Karenin (who would live outside DC) would probably find a different moral code.

For some of the main characters, this book spans the window of their lives where they are showing personal initiative, making decisions that will influence their relationship to Russian society.

Vronsky: is Anna's lover. Essentially no readers like him, and the only characters that do either

think he's hot or fun company in the clubs. He's a young cavalry officer, and one of the first things we hear about him from a minor elderly character are:

"This little fop from Petersburg, they're made by machine, they're all the same sort, they're trash."

Russia has a lot of border disputes in securing its empire, and needs expendable alpha males to manage its army regiments. And Vronsky's regiment does look up to him, and he does purposefully follow the moral code that the culture trained into him--until he decides to make his affair with Anna public.

Vronsky's life was especially fortunate in that he had a code of rules which unquestionably defined everything that ought and ought not to be done...his present relations with Anna and her husband were simple and clear. [They were clearly and precisely defined in the code of rules]...but recently there had appeared new, inner relations between himself and Anna that frightened him with their indefiniteness. [Just yesterday she had announced to him that she was pregnant. And he felt that this news and what she expected of him called for something not wholly defined by his guiding code of rules.]

This frightening indefinite relationship with Anna causes Vronsky, for the first time, to take some actions that the Petersburg rearing did not design him to do. Sex drive is the noise term in Russian aristocracy, which so much of Russian cultural norms work to curate and minimize, but sex drive still produces unwanted surprises in rare individuals like Anna and Vronsky.

I also hated Vronsky at first, but began to feel an uncomfortable kinship with him later in the novel. My rearing has curated me into an intellectual, not an alpha soldier. And Vronsky shifts into a disconcertingly sympathetic character to me.

Here's the narration of Vronsky's troubled solitary thoughts

'Of course,' he said to himself, as if a logical, continuous and clear train of thought had brought him to an unquestionable conclusion. In fact, this 'of course' that he found so convincing was only the consequence of a repetition of exactly the same round of memories and notions that he had already gone through a dozen times within the hour.

Tolstoy is great at making fun of people for thinking they are being courageous or wise when they are really just following everybody else. It's funny until I see myself in his examples, and then I get uncomfortable and less self-assured.

And then, in my first weeks of COVID social distancing, I read about all the ways Anna and Vronsky are passing the time in exile from Russian society.

In between sessions of amateur piano-playing from what I learned in school and lessons, I read

how Vronsky has taken up painting:

He had an ability to understand art and to imitate it faithfully, and thought he had precisely what was needed for an artist. After some hesitation over what kind of painting he would choose--religious, historical, genre, or realistic--he started to paint. He understood all kinds and could be inspired by one or another; but he could not imagine that one could be utterly ignorant of all the kinds of painting and be inspired directly by what was in one's soul, unconcerned whether what one painted belonged to any particular kind. Since he did not know that, and was inspired not directly by life but indirectly by life already embodied in art, he became inspired very quickly and easily, and arrived as quickly and easily at making what he painted look very much like the kind of art he wanted to imitate.

And as I'm considering possible career shifts to things I consider more altruistic, I read how awkward another character felt as Vronsky was proudly throwing his inherited wealth into the construction of a provincial hospital. Meanwhile Anna was patronizing and tutoring some foreign-born orphan, but these were only the causes of later fights, as Vronsky later calls them "unnatural concerns."

And as missing the days of playing sports with friends, I read how a same-generation mother felt about Vronsky playing doubles tennis:

She did not like that general unnaturalness of grown-ups when they play at a children's game by themselves, without children.

All of these quotations made me squirm, and some made me angrily shut the book for a couple weeks.

I'm really troubled by Vronsky, because I can't reassure myself by any rigorous method that he is more shallow than the other characters. Sure, other characters think little of him, but he is a jock (his greatest passion is being a horse jockey), so popular opinion will tend to confirm the stereotype. He does demonstrate independent thought, although he doesn't go through long stretches of self-doubt like some other characters. Maybe time-away-from-self-assuredness is the best metric for shallowness, in which case Vronsky and Karenin are more shallow than Anna and Levin...

At this point, I should mention what you already suspect, that I am a young bachelor. I connected the most with Tolstoy's young bachelors, and least with his women. I feel bad about this, but vis. above there are several ways this novel made me feel worse. I've seen many online reviews, mostly by reviewers with female names, who felt Anna was the star of the show and also loved Kitty. In comparison to what you can find elsewhere, I don't have much worth sharing about the three female leads.

Kitty is the young maiden trying to find her way as an adult, mostly but not exclusively with

regard to her romantic relationships. Unlike most romance novels, the "she got married and lived happily ever after" happens halfway through the book and much of her story arc is about that "ever after."

Anna's sister in law...is the one character where I don't think Tolstoy challenged himself. She is the social conservative mother of the house who has regrets, but ultimately decides that raising children is more fulfilling than fighting society's expectations as Anna does.

IV.

This book is about Russian aristocratic life in the 1870s. Nabakov says as you follow the characters you get an immersive experience of the passage of time as the characters perceive it. One of the ways Tolstoy does this is spend a lot of time with their moment-to-moment thoughts, which involves lots of details of their politics, personal finances, diet, leisure reading, planning for workday and social itineraries. So much of the book is a contrast of how two characters respond to analogous situations. Because this book is such a detailed account of life 150 years ago, it invites comparison from their time to ours. Here are some of the aspects I found the most surprising:

On the day of the horse races, Vronsky came earlier than usual to eat his beefsteak in the common room of the regimental mess. He did not need to maintain himself too strictly, because his weight was exactly the regulation 160lbs; but he also had not to gain any weight, and so he avoided starches and sweets.

Wait, I thought the low-carb beefsteak breakfasts was a new 2018 diet. This translation is from 2000, during the low-fat diet fad. Tolstoy does keep mentioning how Vronsky had great teeth. Also, 160lb seems like a low regulation weight for a cavalry officer who has to be strong and is generally a copious drinker. It is the median accepted weight for a 5'11" male in present-day United States Army. (see US Army enlistment weight requirements) How tall were 1870s Russians? We know the better-fed aristocrats were taller than the commoners. Barbara Tuchman blamed malnutrition when citing how 1900 England lowered their minimum military height requirement from 5'0" to 4'10".

There's also societal declines attributed to recent technology, like how the "fear of missing out" makes people flaky even after they arrive at some other event. You see some adolescent staring at their phone, and then they say "gotta go". This has to be a new thing, right, because how do you leave an event that doesn't have the people you wanted to see if you can't text the people you do want to see? 19th century aristocrats had the footman. There's a scene where Anna's at her socialite friend's house, and sneaks her own postscript into her hostess's letter to the footman so she can meet Vronsky.

V.

The descriptions that make it most clear this novel is of an earlier time are regarding medicine.

Of the three point-of-view women characters, two have near-death childbirth experiences that shock the men around them into character development. The other woman has already had her five children, and we read how in her early 30s her youth and beauty are behind her. Death in childbirth seems a common literary trope. In that era, mothers had at least 0.5% chance of dying in any one childbirth, but now Russia overall has a 0.02% mortality rate. The characters are of the upper class, who burn through money away from home so the women can have the best city doctors during childbirth, so I'm sure the characters would have less than a 0.01% mortality rate today. During the years Tolstoy wrote the novel, three of his children died in infancy.

Then there's the autobiographical sequence of Levin's brother slowly and pathetically dying of tuberculosis. We see him in a German spa, with another character who was diagnosed as potentially tubercolic.

And the famous doctor presented his plan of treatment by Soden waters, the main aim in the prescription of which evidently being that they could do no harm.

We later see him slowly dying in a provincial hotel. His brother and sister-in-law stay with him for 10 days, waiting for him to die, during which time he is alternatively at peace with impending death or miserably suffering from bedsores and anxious to recover, while sometimes "a desire for his death was felt by everyone who saw him." Because he's so provocative, I suspect Tolstoy would have opined on physician-assisted suicide if it existed in any form at that time. These four chapters read a lot like (SSC post) *Who By Very Slow Decay*, so the medical dramas are not completely different today.

This novel took place after the invention of tuberculosis sanatoriums, but shortly before Robert Koch discovered the bacterium that caused the disease, and a lifetime before any effective vaccinations or cures. People still die of tuberculosis, but famous authors and their siblings stopped dying of tuberculosis halfway between *Anna Karenina* and the present day.

VI.

I have to write about the peasants.

Levin's heart is in the country estate, in peasant farming and hunting, but he goes to the city twice on account of a woman: first to court her, second so she can go through childbirth with the best city doctors.

You hear a lot about money in this book, about other character's debts, but also about Levin's spending in the city and country.

The first time he goes to some fancy restaurant to see his friend and learn about his friend's sister-in-law he wants to marry. They eat seafood with a fancy sauce even though Levin would rather eat something simple, and the fancy-sauce meal costs 14 roubles. When he returns to his farm, you learn about the peasants in the course of hearing his springtime dreams for the farm. Levin asks his steward in increasing exasperation why all the work is behind schedule, and the steward says

'We need to hire more workers, but they don't come. There were some today, but they

asked 70 roubles each for the summer.'

Levin kept silent. Again this force opposed him...he could not help fighting even so. He replied 'send to [farther villages] if they don't come. We must look.'

At this point we had already heard Levin's first response to his pathetic, consumptive brother's proposal of Marxism.

[Levin's brother] pointed to some small iron bars tied with string. 'That's the beginning of a new manufacturing association we're undertaking.'

Levin was almost not listening. [...] His brother continued, "you know that capital oppresses the worker, [...] and so we're organizing an association in which all production and profit will be common property."

"Where will this association be located?" Levin asked.

"In the village of V__"

"Why have a metal-working association in a village? I think there's enough to do in the villages without that."

Levin eventually tries this shared-property idea with his farm, and it doesn't work well. But we know from Tolstoy's life that he will eventually forsake the concept of private property and continue to obsess about peasant farm labor until his death. What's striking about this early exchange is Levin's assumption that small-town Russia had particularly more demand for work than supply of workers. This situation is totally flipped 150 years later, where politicians particularly speak of protecting jobs in small-town America and some economists question whether many rural towns should continue to exist. I want to share one of Tolstoy's points in the novel to speculate on part of why that flip happened.

Later in the novel Levin is married, and his in-laws persuade him to take his pregnant wife to Moscow so she doesn't die in childbirth at the hands of some less-competent country doctor. There's a page about Levin's spending in Moscow, how the money starts to fall away from him, and he knows how much the rouble is worth on his underfunded estate but somehow spends so much more in the city.

This is the image Levin provides of inequalities of the time. On the one hand, you see the peasant wages in contrast to the aristocratic spending which is a hot topic today. But there's another economic inequality going on, not of the money to the rich vs the poor, but the money going to the city vs the country. The aristocrats have the money, the dreams, and the agency to craft some part of the world as they want. Levin wants his farm to thrive more than he cares about what the city has to offer in the way of art, cuisine, and all the other things aristocrats purchase to look aristocratic, yet he finds himself spending all his [slack](#)/money on these city things in order to keep up appearances. All the money he saved from not paying for enough peasant labor to make his farm as he desires goes into spending side-effects of his desire for his wife to have the best doctor in her childbirth. One person for two reasons at two times pushes Levin to support the economic forces of the time even as his broader values regret it. Now Levin/Tolstoy is a character of anomalous curmudgeonly intensity, so it's hard to say how

scalable this idea is.

Now, 100 pages earlier, Anna's lover was boasting of how he was building a first-class hospital in the country, one day's carriage-ride away from Levin's estate. It's hard to say how Tolstoy feels about this, because everything Anna and her lover do in social exile is "unnatural."

In case you were considering the utilitarian extreme of charity, the one Jane Addams rejected, this is what and how you should be eating:

The old man crumbled some bread into a bowl, kneaded it with a spoon handle, poured in some water from his whetsone box, cut more bread, sprinkled it with salt, and turned eastward to pray. 'Here master, try a bit of my mash,' he said, squatting down in front of the bowl.

Levin has some of the mash and finds it surprisingly delicious and filling. When he comes home that night he tells his housekeeper and brother he's already eaten dinner and isn't hungry, but they eventually get him to sit down with them, calling upon his responsibilities as a host (is all of Levin's economic consumption about reluctantly keeping up appearances?), and remark about how ravenously he is eating.

VII.

Levin is very much the intellectual ancestor of Taleb and Peterson. When he's feeling down, he lifts weights or swings a scythe. He has this intellectual relationship with his half-brother academic which reads eerily like Taleb's discussions about Nero.

Taleb writes this in Antifragile:

Nero was the victim of an aesthetic ailment that brings revulsion, even phobia, toward: people wearing flip-flops, bankers, politicians...

And this from Levin's opening scene:

Levin was silent, glancing at the unfamiliar faces of his friend's two colleagues and especially at one clerk's hands, with such long white fingers, such long yellow nails curving at the tips, and such huge glittering cuff links on his sleeves, that these hands absorbed all his attention and did not allow him any freedom of thought...Levin looked with hatred at those hands.

Here's Taleb again, about scholarship:

Nero had been spending all his adult life writing a philosophical-technical book. His tendency was to abandon the project every two years and take it up again two years later. He felt that the concept of probability as used was too narrow and incomplete to express the true nature of decisions in the ecology of the real world.

And here's Levin about his book which variously consumes him throughout the novel.

Levin had begun to write a work on farming, the basis of which was that the character of the worker had to be taken as an absolute given in farming, and that consequently, all propositions in the science of farming ought to be deduced not from the givens of soil and climate alone, but also from the known, immutable character of the worker...

[100 pages later] Levin read the politico-economic books...which gave not even the slightest hint of what, he, Levin, and all Russian peasants and landowners were to do with their millions of hands and acres so that they would be most productive for the common good.

[100 pages later] "Oh, I've given up on that book. It's just theory, what's the point?"

[100 pages later] "I'm actually writing a book on agriculture..."

My point in comparing to Taleb is that Tolstoy's Russian peasant obsessions are a case study of more universal questions, namely 1) how to utilize the labor of the masses in a changing economy, 2) how to deal with the mess when theory meets practice, and 3) whether any of these larger questions are more important than one's love life. Tolstoy's answers are 1) this is the most important question a society can ask, 2) this is really hard and he's tried things but hasn't figured it out yet, and 3) no, and if you find yourself wondering about this that's a sign that you are in love and should be doing more about it. The general answers are cliché, but the specifics are interesting.

The strange part about Levin/Tolstoy is that Tolstoy became a prophet of the progressives (Ghandi, Jane Addams, MLK...), and yet Tolstoy was frequently a social conservative who wrote *Anna Karenina* as an argument against the progressive social trends of the day, and Tolstoy the character has the strongest resemblance to Taleb/Peterson. Also, Tolstoy is a strange prophet, because he was excommunicated from the Russian Orthodox church.

VII.

This novel shows a lot of resentment for Tolstoy's on aristocratic class, and it gets more intense in the second half of the novel.

Anna and Vronsky run away with each other, and essentially try to purchase and hobby their way to happiness, and it doesn't work.

Meanwhile, Levin marries Kitty, and then his story reads a lot like an unsuccessful play of the Oregon Trail computer game where you waste all your ammunition shooting squirrels in Nebraska, and then one of your children drinks from the Platte river and dies of dysentery. In fact Levin goes hunting with a couple buddies after Great Snipe, a tiny bird of the Russian swamp. Hunting snipe seems to be the one non-scything-and-childraising activity that Tolstoy gives a pass in the second half of the novel.

And the three gentlemen who go out on a days-long expedition end up eating through all the other food Kitty packed for them and eat all of the snipe in one night. So it's a net-negative

caloric activity, even ignoring Levin's hunting dog who ends up getting more point-of-view sentences than Anna in a 100-page stretch. Then Levin and Kitty go off to the provincial hospital to watch his brother die of TB.

I enjoy reading contrarian essays, but usually the author's surety is off-putting. When Tolstoy writes of his convictions in essays, I usually want to hit him. But I'm more tolerant when he sprinkles his opinions through an autobiographical character in his novel. You see how Levin feels about issues, but simultaneously see when he struggles to articulate his views or is unsure of himself. And you also see how he's an awkward outcast who has spent half of his life alone on a farm because he could not manage to charm any women of his own age and class.

You can start to see how Tolstoy feels about the poor and the wealthy, and how he was perceiving just about everything the wealthy, educated class held dear as pretense, which helps to explain how he became so committed to living like a peasant in his later years.

Shortly after finishing the novel, Tolstoy had a spiritual rebirth and decided that writing novels was a waste of time, and spent the second half of his life writing preachy essays. The second half of this novel in particular shows the seeds of this radicalization. I see Tolstoy the man as a cautionary tale. Don't be too intense, taking ascetic ideas more seriously than anybody around you, or you'll end up a helplessly bitter, difficult man like him.

VIII.

This is a realistic fiction novel with semi-autobiographical elements, so we can summarize the next 150 years of Western society to be an epilogue.

Divorce laws have changed, and there's less stigma than 150 years ago. Maybe Anna Karenina would have survived if she was born to a later culture.

Tolstoy really hates on opera-house music throughout the novel--ideas he develops later in "What is Art," where he claims that the only true music is the stuff peasants sing. This idea makes me angry and uncomfortable. And yet, it's hard to dispute that the most influential music of the century after Tolstoy (the blues, inspiring jazz and rock and more) came out of the farmfields.

At the end of the novel, the characters speak of some Serbian revolt, and share some attitude they've read in the papers that the Russians have a duty to free their fellow Slavs from the yoke of foreign powers. And now the 2010s issue of the annexation of Crimea and the Russian fighting in Ukraine has some broader context. I remember reading several years ago that the Russian TV says they are helping the Ukrainians who identify as Russian, and thinking that was ridiculous, but now seeing that attitude as part of a centuries-long myth in Russian culture.

And we can follow the story of Levin and Kitty by reading the biography of Tolstoy and his wife. They stayed married, but it wasn't happy at the end. Tolstoy wanted to renounce his novels and

spread all his inheritance to the peasants. His wife disagreed about this. It's hard to say what role jealousy played here. Tolstoy fathered a child with a peasant, as he recorded in a diary he gave to his fiancée the night before the wedding, and he never kicked the mother or bastard child off the land. I can imagine the conversation: "The inheritance should go to all the peasants."
"Oh, did you have any particular ones in mind?"

Then seven years after Tolstoy's death came the Bolshevik Revolution, and his family fled to Sweden :)

P.S.

Most of you will understandably choose to never read Tolstoy. If you are suddenly feeling inspired to read everything I've left out, I should mention a couple things. If you are reading any remotely literary foreign book for enjoyment, the translator matters, and Pevear-Volokhonsky are the best Russian-to-English translators. This is popular opinion. I used to think opinions about translators was pretension, then I read two Dostoyevsky books by two different translators. I preferred the plot and characters of the first book, but enjoyed the writing style and conversations in the second one more. Then the thoughts and conversations of this Pevear-Volokhonsky translation of Anna Karenina felt even more natural and helped me to empathize with the characters. The original author is still a confounder in my n=3 experience, but I hadn't heard the opinion about Pevear-Volokhonsky until after I started reading their translation. Pevear-Volokhonsky write about how extensively Tolstoy brought nuance and irony with his use of Russian and how that makes him difficult to translate, which makes it mystifying why Tolstoy championed Esperanto.

Apocalypse Never by Michael Shellenberger

I

Here's Michael Shellenberger's solution to our environmental problems: let's all get rich.

But aren't all the cool people on Twitter talking about degrowth? They are, and they are wrong.

Rich countries have a better record on many issues specifically because they can afford this. Sure, they are responsible for more carbon emissions, but it's not because people in Congo choose to lead a clean lifestyle. They would love to have a steady power source (even if it would burn coal), and we should thank them more often. Also, rich countries are resilient. Climate change may cause more extreme events, but hurricanes and fires are less dangerous when you have money to build better and evacuate faster.

But can't we talk about climate as an existential risk even if it's not? Maybe Extinction Rebellion is wrong about humans dying out in ten years, but what's so bad about getting a wonderful green world of solar and wind even if we didn't really need it? Argument #1: [One in five UK kids have nightmares about climate change](#). Childhood may be less pleasant if people around you are constantly talking about the end of the world in 2030. Argument #2 (which, it seems to me, isn't explicitly made by Shellenberger): some problems are better dealt with in non-obvious ways that require longer timescales. Argument #3: people just give up.

Shellenberger thinks that some environmentalists have misled us into alarmism. Here are some of the stories where they have been wrong:

1. Fires in the Amazon rainforests don't kill the planet's lungs, because this isn't how the lungs work (forests produce *and* consume oxygen, so the net contribution is effectively zero). Deforestation is a problem, but not a new one—people are worried about Brazil converting large areas into pastures and soy fields, yet European farmers won't be ordered to plant the trees and leave their lands. Greenpeace and other NGOs pressured the Brazilian government into

unnecessarily strict forest laws, and this led to Bolsonaro's presidency. Also, Greenpeace conflated Cerrado savannahs with Amazon rainforests; World Bank published reports saying that modern intensive agriculture is bad; Macron revitalised the deforestation debate to protect French farmers.

2. Plastic straws do kill turtles and plastic recycling basically doesn't exist (dumping all your stuff to China or Malaysia doesn't count as recycling). But there are [some studies](#) showing that maybe plastics can disintegrate faster and better than previously thought (thanks to sunlight and microbes). More importantly, even though plastic stuff is killing a lot of marine life now, humans killed many more elephants and turtles before they invented plastics. Even more importantly, plastic stuff is not as devastating as fishing boats are which kill both directly and through overfishing.
3. Animals are dying out, but we are far from a global extinction event. That popular book you've seen relies on a specific model too much, and this model isn't doing great. Also, we are pretty good at saving very small populations and we are already protecting about 15% of the land surface as important habitats.
4. The whales have been saved not by activism, but by capitalism. By the time International Whaling Commission created a moratorium in 1982, 99% of all whales killed in the 20th century had already died; also, IWC set quotas too high. Whales have been saved by kerosene and palm oil becoming cheaper, not by Greenpeace. Notably, countries without free markets have continued whaling for longer (not just USSR but also Japan and Norway that had protectionist policies).
5. Carbon emissions from meat are noticeable, but cutting just those won't be enough. Even if we all went vegetarian or vegan, total carbon emissions would go down by five, maybe ten percent. Free-range farming sounds great, but pasture beef requires fifteen times more land and emits three times more carbon than industrial beef (also, [cows like predictability and cleanliness](#), so they might be okay with not roaming the fields).

There are more, obviously; I'm leaving out a lot of IPCC reports, penguins, and fires. And many of these stories are definitely simplifying difficult question from almost intrinsically hard fields. Modelling climate change is definitely hard. Measuring plastic levels in seas takes a lot of time. Even counting animals requires either many high-

quality cameras and computer vision advances or an integrated solution that's also known as a human sitting and watching a single tree for twelve hours a day.

(And after you've counted some animals, you're left with questions like what to do with animals you can't count because they died out eons—or fifty years—ago. [Here's a paper saying that we are in the sixth extinction](#) because current extinction for species is much higher than the standard rate between the mass extinction events. But how sure are we about species that died out in the 1600s? IUCN Red List (source for the Threatened-Vulnerable-Extinct plaque in Wikipedia) [has a graph](#) showing that more assessed species get classified as threatened each year, but the percentage of threatened is going down. Should we even be counting species since they are so dependent on whether we have enough biologists to classify everything?)

II

What do all those stories have in common? Weirdly, Shellenberger isn't very interested in why wrong versions have become popular or what are general working solutions. Instead, he is very interested in talking about Malthus.

A lot of people in the twentieth century were worried about overpopulation. Some of those people (including Paul Ehrlich—the worse Paul Ehrlich, of course) also thought that [“giving society cheap and abundant energy would be the moral equivalent of giving an idiot child a machine gun”](#) (apparently, because this energy would make the humanity to industrialise every corner of the planet; also, he thought we should all adopt ways of living that are traditional in French Polynesia. I'm sorry for talking about Paul Ehrlich in polite society.)

Of course, overpopulation didn't happen. But this thinking led to rich countries and global organisations putting their aid money in charity and not in infrastructure development. If we are not helping poorer countries, let's at least not stop their attempts at industrial growth, says Shellenberger.

The current degrowth vogue probably also partly stems from these Malthusian roots. But even more important is the fact that most of those calling for stopping growth don't know enough about poor countries. We can't just redistribute everything—not because

it's politically, but [because it isn't enough](#). There will be a lot of growth in poor countries, and it will lead to more carbon emissions—and an enormous increase in human well-being.

If people are indeed actively fighting against power stations in Africa, we should fight back. If people are started to burn gas instead of coal, we shouldn't immediately think they've sold their souls to Gazprom. But I think we should be interested in why people worry about straws instead of overfishing. "We humans should take as little space as possible" is a useful takeaway from Shellenberger's energy density discussions, but it doesn't help in every case.

Some issues make people rise up to immediately call for a ban on this barbaric and medieval thing that can't be tolerated in modern society. Free-range farming sounds great because of the beautiful image of a cow running its first non-steeplechase mile. However, banning industrial farming will hit not only the poorest people but also the wildlife for which you don't have an immediate image. Non-ban decisions need explanations and transparency, but people can like cute non-intuitive explanations ([sometimes even too much](#))!

And when you do need a ban, just redeploy Extinction Rebellion.

III

Shellenberger spends a lot of time on the ways solar and wind aren't enough because of the unsolved storage issues—batteries aren't there yet, and [pumped-storage hydroelectricity](#) needs tenfold more dams in the US to handle everything (pumped hydro is putting the water up a hill when it's sunny or windy, and letting it go down to spin the generators later).

(I think Shellenberger is attacking a straw man here. Sure, that particular plan of putting everything into dams isn't going to work, but we don't need to do it. Real Engineering has [a recent video](#) on pumped hydro in Ireland (including a visit to [the Turlough Hill station](#)). His estimate is that Ireland would need about 37 such stations to switch from gas to wind completely. Building them is expensive (about one billion dollars per station) and lengthy; more importantly, you need some help from geography (ideally, two

freshwater lakes, one on the hill, another just below that hill). In some places you can use seawater, and there are ideas about putting water underground, but it all won't happen any time soon. In practice, however, you need all 37—Li-ion will get better, hydrogen will get better, some efficiencies will increase, and helping all that with a station or ten might be enough.)

(On a related note, [existing dams aren't in a good shape](#), and we should spend a lot on their maintenance to avoid [another Banqiao](#).)

More importantly, both solar and wind need a lot of space, and that space is being taken away from wildlife (noise is bad, mercury is dangerous, getting your wings crushed by a fan is lethal). Meanwhile, nuclear plants are comparatively small, provide a steady baseline and emit zero carbon. Also, they are extremely safe.

And despite all that, people are afraid of nuclear plants. How did that happen?

Shellenberger's non-obvious answer is the Brown family famous for being disproportionately represented among California governors.

In the late 1960s, the Indonesian state-owned oil company Pertamina asked Pat Brown, who just left the governor's post, to help them raise some money on Wall Street. Brown raised some 13 billion and in return got a concession to sell Pertamina's oil in the US and to lobby for its interests in Washington. Quite soon California enacted a regulation that would allow only low sulphur oil to be used there—and, would you believe it, that's just the type of oil found in Indonesia (an attempt to build a refinery for Alaska oil has been blocked).

Pat's son Jerry, both as governor between 1975 and 1983, and in his later career was very keen on closing existing nuclear plants in California and killing any plans for the new ones. Sure, he didn't own anything himself, but his sister Kathleen was on the board of the company responsible for [a massive gas leak](#) that might have been ignored for too long by the government of California.

Shellenberger has more such stories about lobbying and hidden interests: for instance, if you're Tesla and you own SolarCity, why wouldn't you [tell the government it's a good](#)

[idea to kill the last remaining nuclear plant in California](#)? If you're a traditional fossil fuel company, you love to be seen near wind and solar, and those still need your gas baseline, so why wouldn't you put Greta Thunberg pictures in all your corporately sincere #solar tweets while spending your actual efforts on lobbying against nuclear?

Other examples are less convincing. I don't think you can call Michael Bloomberg a "major investor in natural gas" based on [an article](#) in which the main evidence is Bloomberg's hedge fund manager saying "yeah, maybe we will invest in gas a bit more". Should you chastise Al Gore for selling a TV channel he co-founded to Al Jazeera (which means Qatar which means oil)? I don't find this purism helpful: Sierra Club taking money from a gas company CEO and coming out against nuclear is evidence of dishonesty; having a small percentage of your money in stocks of gas companies isn't.

Overall, I think Shellenberger overestimates the politicians' role in the decline of nuclear power. He has a scene where Jerry Brown comes to an approving crowd at a "[No Nukes](#)" concert in 1979, but people weren't there for Brown. Shellenberger talks a bit about [Jane Fonda](#) and Simpsons but is more interested in journalistic investigations. Fair enough, it's not a media studies book. And yet, hundreds of millions more people have seen the nuclear plant in Springfield than have heard about Jerry Brown (hundreds of millions would be surprised to learn that Arnold Schwarzenegger isn't a governor anymore). I don't think I have a complete understanding of why Chernobyl is a household name and an HBO series but Banqiao and [Bhopal](#) aren't. It is mostly a product of confusion caused by nuclear bombs using the same uranium, plutonium and the word "nuclear"?

But no matter what happened previously, it is now on nuclear plants builders and proponents to fight the misconceptions. I don't think we are doing a good enough job here.

IVa

From 2012 to approximately 2021 Belarus has been building the [Astraviec nuclear power plant](#), and by "Belarus has been building" I mean "Rosatom has been building in Belarus using money provided by the Russian credit". Rosatom is a great company, and

[reactors of the same type](#) are either already built or under construction in Russia, [China](#), [Egypt](#), [Bangladesh](#) and [Finland](#). Despite this, Belarusian nuclear has been a disaster.

Of course, there's nothing wrong with the reactor itself. Belarus isn't sunny or windy, but it doesn't have serious storms or earthquakes, so building a nuclear plant here is a good idea. You just have to convince the population scarred by Chernobyl that what you're is fine and great and useful, and the way to do that is definitely not through empty repetitions of "everything's fine, you don't need to know more" mantra. It doesn't matter whether a random construction element falling is a serious mishap, what matters is whether people think it is. "Something happened, and the government doesn't want to talk about it" is an instant pattern-match for Chernobyl. You can't fight it by putting out interviews with mumbling bureaucrats who look like just came back to Belarus after receiving a medal from Brezhnev. You can only fight it with honesty and good explanations.

In Minsk, I've seen several billboards praising wind as a source of energy (our ancestors used it!), but not a single one about nuclear safety. I don't think Belarusian state TV channels have even tried making a good documentary explaining how Astraviec is different from Chernobyl (one can find some token attempts on Youtube, but they are either [badly made by a local TV station](#) or [badly made by Rosatom itself](#)). If you're spending 10 billion to build the plant, you shouldn't worry about spending 10 million to build great messaging about it.

Of course, it's all completely moot for some time. Belarus is now a failed state, with its government being laser-focused on making the word "ridiculous" unusable ([10 years for being a witness of a murder by a policeman? 4 years for throwing a flower into riot police's direction?](#) Talking about creating own vaccine because you spent too much time denying COVID, and anyway you need money to pay the military to protect you?). There aren't any TV ads explaining how the nuclear plant works, because there are only ads praising the army and the riot police. Still, there will be a new government, and nuclear in Belarus deserves a better future. In the end, the old government's not doing any propaganda work around it might even be helpful.

IVb

Smog is the worst thing about living in Kraków. People have been successfully selling masks in Poland before 2020. [Scientists in the Kraków's Jagiellonian University are the world's leading experts the air pollution allergy](#). I thought about buying noise-cancelling headphones just to battle the second air purifier in my apartment.

People are telling me it's been worse. Kraków has now banned burning any wood and coal inside the city by private citizens. Poland is now [generating less than half of its electricity from coal](#) (although it's still 46 percent, and it's been replaced by oil and gas; it also probably happened mostly because [Poland imports more electricity by the year](#), since burning Polish coal is becoming too expensive).

Aside from just buying power from outside, there are plans for more renewables (wind) and (finally) nuclear (six (!) plants with the first one up and running by 2033). Poland had a failed attempt at nuclear once before: [Żarnowiec](#) nuclear plant has been started in the 1980s but was abandoned in 1990 after Chernobyl-inspired protests and a referendum in the Gdańsk voivodeship. Some parts of Żarnowiec are now in use in the Paks nuclear plant that generates more than 50 percent of Hungary's electricity.

I'm hopeful but not entirely optimistic. On the hand, it seems too easy to make fun of [a party that thinks LGBT is the main threat](#) to Poland's future and instead of solving an existing crisis (COVID) creates a new one by banning abortions at the worst possible moment. On the other hand, there will be protests both by misguided environmentalists and [redundant coal miners](#), and it might be too nice and easy for any party to cave in and delay this mess even further. Luckily gas is too Russian for anyone's liking here.

And there is already a lot of complacency from the government's side. There won't be any building until 2024. Nuclear plants take this long and cost this much partly because they are large and complicated machines, but also because we decided to have much stricter safety requirements for them (compared to, say, chemistry factories, and coal mines). But you're the government! You can change it! (Maybe Brussels won't like it, but Brussels didn't like killing independent judiciary and media, and yet here we are.)

V

Shellenberger doesn't have a chapter full of ambitious yet completely realistic solutions. Instead, he spends much of the last chapter arguing that environmentalism is a religion. This seems to me if not exactly something in a bad faith, but at least completely useless. Maybe environmentalism does check all the requirements from your favourite definition. Maybe now you feel better about yourself when you can't win an argument against an anti-nuclear alarmist. But weren't we talking about how we need to make this whole environment stuff more about facts and less about feelings?

No, says Shellenberger and decides to go full Haidt against apocalyptic environmentalists. They have spirituality and death-adjacent excitement? Let's one-up them! And... and that's it going to happen? "<..> we need to go beyond rationalism and re-embrace humanism, which affirms humankind's specialness, against Malthusian and apocalyptic environmentalists who condemn human civilization and humanity itself. <..> we must ground ourselves first in our commitment to the transcendent moral purpose of universal human flourishing and environmental progress, and then in rationality."

I don't know. It seems to be super-standard consequentialism? But with environmental progress, which just converts to higher weights for non-humans?

A bit further, Shellenberger starts his epilogue with the sentence "Few things make one feel more immortal than saving the life of a nuclear plant." People who are protesting the closures of coal mines that existed for hundreds of years probably feel the same. Let's just save nuclear plants because they are small, efficient, and clean. One day they will be replaced by fusion and better batteries, and I hope Shellenberger won't feel a need to protect a fission plant.

I don't think we should get into this game of "who can better exploit human moral foundations". On [a recent Rationally Speaking episode](#), Haidt was still talking about the left not utilising all these foundations completely, but it seems to me that all people are getting better at this. Supporting solar and wind and hating nuclear is right up the purity alley, and loyalty importance on the left might have been underestimated by Haidt.

Understanding why people don't like nuclear is useful. Finding out why apocalypse pronouncements are so popular may help us. And there is nothing more important than development economics. But we don't need to make a religion out of it. Esther Duflo can be wrong, Saint Esther can't. You can tell inspiring stories of progress and showcase examples of growth helping people (as Shellenberger does) without sliding into an emotional contest. You can explain things to people. You can tell people something is x when it's x , even if $10x$ would make them care more about this problem.

You don't need to tell me I'm a naive optimist. I know that.

Astounding: John W. Campbell, Isaac Asimov,
Robert A. Heinlein, L. Ron Hubbard, and the Golden
Age of Science Fiction by Alec Nevala-Lee

Every now and then, amid your fevered cries for net neutrality, free soil and free silver, the restoration of the house of Stuart, more episodes of *Firefly*, or whatever other hopeless cause gets your blood racing and your family members fleeing (they recognize a wind-up to a full-fledged rant when they hear one), against all odds the universe actually hears, takes note, and gives you precisely what you've asked for - not often, dammit, but sometimes.

Thus it was that after decades of buttonholing strangers and lecturing them on the nation's desperate need for a biography of John W. Campbell, the pioneering science fiction writer and influential editor of *Astounding Science Fiction* (later *Analog*) from 1937 until his death in 1971, a couple of months ago I discovered that just such a book had finally been written.

I Immediately put *Astounding: John W. Campbell, Isaac Asimov, Robert A. Heinlein, L. Ron Hubbard, and the Golden Age of Science Fiction* by Alec Nevala-Lee at the top of my Christmas list, and I have just finished devouring it, blurbs, book jacket, binding glue, and all. Give me a second to belch, and I'll tell you what I thought.

As the subtitle indicates, Nevala-Lee's book is an even bigger feast than I was hoping for. *Astounding* is nothing less than a quadruple biography, chronicling not only Campbell's life, but also the lives of his three most significant collaborators; it also highlights the substantial contributions the four men's various wives made to their husbands' careers. (I think there were eight or ten wives, all told – it was hard to keep track. Being married to these fellows was the matrimonial equivalent of the blood test scene in John Carpenter's film, *The Thing*, which was of course based on Campbell's

classic story of a shapeshifting alien, “Who Goes There?”) Nevala-Lee rightly sees the saga of science fiction’s Golden Age as a story of the collective effort and shared vision of a bunch of rugged individualists, which is a neat bit of irony.

There's a wealth of surprising information here, and I guarantee that even if you consider yourself well-versed in the subject, you're going to find out a lot that you didn't know. (Some of it, especially in the chapters that cover the birth and growth of Dianetics, are things that you might rather *not* know.) For example, it's common knowledge that Cleve Cartmill's 1944 story “Deadline” so accurately predicted some aspects of what was then going on in the Manhattan Project that it attracted the attention both of the physicists working in Los Alamos (many of whom were avid readers of *Astounding*) and of the Counterintelligence Corps, who sent agents to grill Campbell and Cartmill. What was new to me was how Campbell deliberately provoked the incident (the story idea was Campbell's, and he practically dictated the technical details to Cartmill), all to verify his suspicions that the government was constructing an atomic bomb; he actually *wanted* to attract the security service's attention. He never realized how close the whole thing came to getting *Astounding* closed down. It's a typical example of Campbell's blend of far-seeing brilliance and blind recklessness. Campbell rates this kind of serious biographical treatment because he was one of the most significant cultural figures of the twentieth century, and under his editorship, *Astounding* was the single most important and influential venue for serious science fiction that has ever existed in American publishing. Campbell played a greater role in shaping the genre than any other figure; from the time he took over editorship of the magazine in the late 30's until at least the early 50's, he was in effect the Pope of

Science Fiction, the man who defined what the genre was and what it could - and couldn't - do. As time went on, this very dominance caused many writers to chafe under the limits imposed by Campbell's exacting – and sometimes idiosyncratic – requirements.

This resistance ultimately led to a very fruitful reaction in the pages of, among others, *The Magazine of Fantasy and Science Fiction* and *Galaxy*, a reaction that culminated in the so-called “new wave” of the sixties – but even that movement couldn't completely eclipse Campbell or cancel out his authority. Nevala-Lee relates an anecdote about that walking embodiment of rude reaction, Harlan Ellison (who Campbell called “an insulting little squirt with a nasty tongue”) sneering at Campbell as an ossified relic but then being ecstatic when a story he co-wrote with Ben Bova was accepted by the editor for *Analog*.

The seeming contradiction of Ellison's rebelling against Campbell while also desiring to produce something that he would find acceptable testifies to an inescapable fact: to the end, John W. Campbell remained a towering figure in the genre, and for a very simple reason – he was an indisputably great editor, a man who, even in his decline, stood head and shoulders above all others in his field. On the one hand, he was a fountain of the kind of practical, nuts and bolts advice that all editors dispense; he once told Asimov, “When you have trouble with the beginning of a story, that is because you're starting in the wrong place, and almost certainly too soon. Pick out a later point in the story and begin again.”

But in addition to this kind of purely technical advice, he was uniquely suited for science fiction in that he tossed off uncommonly fruitful story ideas with reckless abandon, all for

his writers to make use of, and use them they did. Though he tried to give the credit for the three laws of robotics to Isaac Asimov, Asimov himself always insisted that they were Campbell's creation, along with the basic ideas for "Nightfall" (often cited as the greatest science fiction story ever written) and the *Foundation* trilogy. And that's just a few examples from the career of one writer; Campbell was just as generous with everyone who wrote for him, though not everyone could thrive under this kind of tutelage. (Larry Niven – who you would think would be a natural for *Analog* – didn't write much for Campbell, saying "He liked his ideas better than he liked mine.") A man who was so opinionated and so sure of himself could certainly be frustrating to deal with. Asimov, who never stopped admiring him, described Campbell this way:

"Suppose you meet a man who asks you what your field of endeavor is and you tell him that you are the world's greatest living vertebrate paleontologist, which is, of course, what you are. And suppose that, on hearing this, the man you meet fixes you with a glittering eye and proceeds to lecture you for five hours on vertebrate paleontology, getting all his facts wrong, yet somehow leaving you unable to argue them. You will then have met Campbell."

One of the most prominent aspects of Campbell's character was his need to shake things up, and it was this iconoclastic tendency that was both his greatest strength and his most glaring weakness. It led him to push his genre far beyond the rigid, Gernsbackian boundaries that were the norm when he broke into science fiction as a writer (of precisely those same slide-rule stiff, light year-leaping yarns that as an

editor he moved the genre away from). It was a huge accomplishment that all science fiction readers and writers should be forever grateful for.

But Campbell's reflexive gadflyism also led him into his most egregious follies. On the "more embarrassing than harmful" side, it was an impulse which often manifested itself in an advocacy of various crackpot inventions. (He had trained as a physicist and had just enough technical knowledge to be dangerous.) Taking up space in *Astounding* to sing the praises of the Dean Drive (a supposedly "reactionless" space drive that no one, including its inventor, could ever get to actually work) was silly enough, but Campbell's delight in attacking accepted truths also led him into some highly dubious areas, to say the least, as in his oft-repeated assertion that in some situations, slavery was an acceptable – and even beneficial – social system. How much of this was racism and how much a transgressive delight in shocking the hell out people is impossible to precisely determine. Nevala-Lee doesn't weight the balance in either direction; he just presents the evidence and respects his readers enough to let them draw their own conclusions, a refreshing and increasingly rare quality these days.

Nevala-Lee takes the same approach when it comes to the four men's sexual transgressions; he neither conceals nor excuses anything, but he also doesn't imply that such actions invalidate his subjects' accomplishments or amount to the sum of their characters. This evenhanded approach is seen most clearly in the book's portrait of Asimov, who was a notorious groper, but who nevertheless comes across as the most sympathetic figure in the book.

The most gripping chapters of *Astounding* are those which deal with the early days of Dianetics. Campbell enthusiastically hailed this wild concoction as an exact

science of the mind, which was something he had been seeking for most of his life. Campbell never completely renounced Dianetics, even as it morphed into Scientology and he became estranged from its increasingly unhinged creator, L. Ron Hubbard. It's perhaps not surprising that Campbell, with his fascination with psi powers and his belief in humanity as an ongoing evolutionary project (with science fiction serving as evolution's indispensable handmaiden) was attracted by some of the ideas embodied in Hubbard's work.

What *is* surprising is how easily Campbell and so many of his self-styled group of "competent men," all of whom saw themselves as cool rationalists and objective minds, were taken in by Hubbard's ceaseless torrent of brazen bullshit. Some, like Jack Williamson, instantly pegged Hubbard as a pathological liar, but an amazing number of highly intelligent people followed him like lemmings off a cliff. A.E. Van Vogt – to name just one science fiction writer among many - was an early disciple of Dianetics, and while Heinlein stayed clear of the movement, he always passionately defended Hubbard as a genuine war hero. In reality Hubbard's constant insubordination and flagrant incompetence got him quickly relieved from every command he held during the war, and the limp that Hubbard blamed on Japanese shrapnel was actually the result of falling off a ladder.

The whole thing makes for a full, fascinating story. If *Astounding* has any deficiency, it's that in focusing on the biographical comings and goings of his four main protagonists and the countless people in their orbit, Nevala-Lee has little room left for detailing how Campbell's extraordinary blend of prejudices, obsessions, instincts, and insights concretely shaped the genre, both in the pages of *Astounding/Analog* and in all

the varieties of “anti-Campbellian” reaction, some of which are flourishing even today, almost a half century after his death.

I could have used more in-depth analysis of characteristic works from Campbell’s crew; this kind of critical approach would have put a little more flesh on the bones of Nevala-Lee’s argument. The critical assessments that *are* present are generally extremely brief and the judgments are conventional when they’re not just odd. (Eric Frank Russell’s *Sinister Barrier* is an important work in the development of the genre and remains an enjoyable read, but I’m not sure that “it still deserves to be ranked as one of the greatest science fiction novels ever written.”)

But these are minor quibbles. Though I feel that the definitive account of Campbell and his era has yet to be written, *Astounding: John W. Campbell, Isaac Asimov, Robert A. Heinlein, L. Ron Hubbard, and the Golden Age of Science Fiction* is an excellent start. In the book’s afterward, Nevala-Lee says that he hopes his book will be “a necessary first step in any comprehensive reckoning.” It is certainly that and more; someone has at last laid a solid foundation on which the rest of the history of modern science fiction can be built, and we owe Alec Nevala-Lee our thanks for at last taking on that hard and long-deferred task, and for doing it so well.

Atlas Shrugged by Ayn Rand

[DP]

Allow me to confess: I never intended to read this book, and I knew almost nothing about Ayn Rand beforehand. I had read Yudkowsky’s [Guardians of Ayn Rand](#), which was not flattering to Rand, and later I heard someone say they could tell Atlas Shrugged was “garbage” after three

pages. I was curious what it was that didn't impress people and decided to find out. However, after three pages it didn't seem like garbage at all, and I ended up reading long enough to want to learn the book's mysteries, such as why businessmen kept disappearing, and what was causing the deterioration of the world. I was downloading it one chapter at a time from a random website, in such a way that there was no way to tell how long it was in advance.

Atlas Shrugged, published in 1957, turned out to be longer than I could possibly have imagined: 1168 pages. It's about 3.10 megabytes of plain ASCII and more than half a million words. It is 2.7 times as long as the Arthur C. Clark trilogy "2001: A Space Odyssey", "2010: Odyssey Two" and "2061: Odyssey Three" (these are 1.13 MB in total). It's the kind of book that invites comparisons with the distance to the moon, and if I wrote like Rand does, I would sprinkle nine such comparisons throughout this review. The book is divided into three parts of ten chapters each, but it is certainly not a trilogy, since important mysteries are stretched out from nearly the beginning until nearly the end, like the men who fought so hard over a penny that they invented copper wire.

Reportedly, it sold 29 million copies by 2013, with almost 3 million copies purchased for distribution to schools by the Ayn Rand institute. In 1991, Atlas Shrugged was number two behind the Bible in a survey asking for "the most influential book in the respondent's life". (No percentage is given and I don't have more recent data.)

I spent over 80 hours reading this, and if I knew its size in advance, I would not have bothered. Seeing Scott's call for book reviews is the only thing that made me read all the way to the end, though I couldn't help skipping sentences and paragraphs sometimes.

This review has minor spoilers. Also, I decided that no one else should have to go through what I did, so I prepared a [condensed version of the story](#) for this review on a separate page, because I don't think reading it is worth dozens of hours of your time — but it might be worth two, if only for the value of pretending you have read an entire book by Ayn Rand. I think the story could have fit in one-fifth the word count and still have told the same story equally well.

While Ayn Rand was influential in the development of modern American libertarianism, she rejected the label herself. This quote [from Wikipedia](#) doesn't explain why to my satisfaction:

All kinds of people today call themselves "libertarians," especially something calling itself the New Right, which consists of hippies who are anarchists instead of leftist collectivists; but anarchists are collectivists. Capitalism is the one system that requires absolute objective law, yet libertarians combine capitalism and anarchism. That's worse than anything the New Left has proposed. It's a mockery of philosophy and ideology. They sling slogans and try to ride on two bandwagons. They want to be hippies, but don't want to preach collectivism because those jobs are already taken. But anarchism is a logical outgrowth of the anti-intellectual side of collectivism. I could deal with a Marxist with a greater chance of reaching some kind of understanding, and with much greater respect. Anarchists are the scum of the intellectual world of the Left, which has given them up. So the Right picks up another leftist discard. That's the libertarian movement.

Are what she calls "libertarians" the same as a modern hardcore "all taxes are theft" libertarians? That's not how it sounds to me.

The book is set in some indeterminate future time, “centuries” after the industrial revolution and “centuries” after the fictional Nathaniel “Nat” Taggart built a transcontinental railway across the United States. Taggart built his railroad with no government handouts, and proudly held his goal to be making a profit.

Despite the passage of time, Rand’s future world looks as it would if all human innovation had stopped on a dime in 1957. Wired telephones, telegraphs, radio broadcasts, television and newspapers are the only major forms of communication; typewriters (with carbon copies!) are the primary means of writing; trains are powered mainly by coal but often by deisel; passenger trains are preferred over airplanes; computers don’t exist; prices are plausible for 1957; and many skyscrapers are beginning to crack and crumble from old age and neglect.

He enjoyed the sight of a prosperous street; not more than every fourth one of the stores was out of business, its windows dark and empty. - Chapter 1

The book never once mentions communism, socialism or collectivism, but aside from the United States, most of the countries mentioned are “people’s states” and the others are about to turn into them: The People’s State of Mexico, The People’s State of Norway, The People’s State of France, the People’s State of England, the People’s State of Portugal, the People’s State of Turkey, the People’s State of China, the People’s State of Germany, the People’s State of Chile, the People’s State of India, and the People’s State of Guatemala. In Part 3 Chapter 4, Argentina becomes a People’s State.

Keep in mind that the book was published after many years of McCarthyism, at the tail end of the Second Red Scare, which I’m told lasted from the late 1940s through the 1950s. Rand clearly thought that communism was a serious threat, despite already being fairly unpopular among Americans (an unpopularity she [promoted eagerly](#)). The book eventually gets around to spending ample time attacking communist ideas, but what she attacks is her own interpretation of its most superficial elements, like the phrase “from each according to his ability, to each according to his need.”

Quick reviews from 1957

It is probably the worst piece of large fiction written since Miss Rand’s equally weighty “The Fountainhead.” - Robert R. Kirsch LA Times, 1957

[...] a book every businessman should hug to his breast [...] But how the shabby little left-wingers are going to hate it! - Paul Jordan-Smith, Los Angeles Times, 1957

Gigantic, relentless, often fantastic, this book is definitely not one to be swallowed whole. Throughout its 1,168 pages, Miss Rand never cracks a smile. Conversations deteriorate into monologues as one character after another laboriously declaims his set of values. - Newsweek, 1957

There is much good sense in this book and it deserves more careful consideration than it is likely to get. [...] The worst thing in her book is her denunciation of what she calls mysticism [...] No, Miss Rand, a mystic is a man who insists upon using

those areas of his mind which you block off. - Edward Wagenknecht, Chicago Daily Tribune, 1957

Miss Rand's villains resemble no one I have ever encountered, and I finally decided to call them "liberals," chiefly because I can't imagine whom else she might have in mind. [...] America is plunged into a catastrophic depression, caused by the government's infernal meddling with the economy, and most of the other nations of the world have become People's States, whose inhabitants are actually grubbing up roots to keep themselves alive. The last sparks of industrial competence are concentrated in the minds of two dozen — at most — American businessmen, who manage to hold the globe aloft in spite of the best efforts of governments everywhere to bring it down. - Donald Malcolm, the New Yorker

It runs 1,168 pages, and you won't want to miss one word. [...] You'll say it can't happen here — but it's happening every day and we sit still while watching our rights as humans being whittled away. - Hedda Hopper

[Source](#)

Characters

This book is mostly about the characters in it; the world and the plot are incidental in comparison. Here are the main characters in the first half of the book:

1. Dagny Taggart, Vice-President in Charge of Operations for Taggart Transcontinental, and a descendant of Nat Taggart. She loves railroads more than anyone, because they act as the arteries and blood cells of the country. Her dream in life is to run Taggart Transcontinental.
2. Francisco d'Anconia, a descendant of Sebastian d'Anconia, who built a copper mining empire centuries ago. Each generation of d'Anconia grew the empire slightly larger than the last, but Francisco d'Anconia was a hard-working genius who managed to grow the business far more quickly than his ancestors. But Francisco's behavior changed about ten years ago:
At the age of twenty-three, when he inherited his fortune, Francisco d'Anconia had been famous as the copper king of the world. Now, at thirty-six, he was famous as the richest man and the most spectacularly worthless playboy on earth. He was the last descendant of one of the noblest families of Argentina. He owned cattle ranches, coffee plantations and most of the copper mines of Chile. He owned half of South America and sundry mines scattered through the United States as small change.
"the rebirth of [...] the world — has to start here, in the United States. This country was the only country in history born [...] as a rational product of man's mind." - Francisco d'Anconia
3. Henry "Hank" Rearden is Ayn Rand's new-money hero, who built the most important steel business in the United States on the outskirts of Philadelphia. His company, Rearden Steel, is supplied by Rearden Ore, Rearden Coal, and Rearden Limestone. Rand portrays Rearden as a lone genius (a concept I regard as a [dangerously disheartening myth](#)) who develops an incredible alternative to steel that contains copper and is "half the weight", "twice as safe", and "will last three times longer than any other

[metal]", plus it's good for building anything from chicken wire to railways to microphones to airplanes.

4. James "Jim" Taggart, President of Taggart Transcontinental, is Dagny Taggart's brother. At first, James' primary skills appear to be refusing to make decisions, deflecting, and making illogical statements. We later learn that the board elected him president "in the same manner as they refused to walk under a ladder, to propitiate the same kind of fear. They talked about his gift of 'making railroads popular,' his 'good press,' his 'Washington ability.'" James is a deft influence-peddler, friends with all kinds of Washington elites. James also hates competition almost as much as he seems to hate himself, and is angry that oilman Ellis Wyatt "double-crossed" Taggart Transcontinental (who ran one tank train per week for him on their old track) by doing business with Phoenix-Durango instead (who lets him run two tank trains every day).

Aside from these main characters there are many secondary ones, such as:

- Eddie Willers, who was a child friend with Dagny and Francisco, who used to all hang out together for one month each summer. Eddie is Dagny's competent assistant, basically a more naive and modestly less successful version of Dagny.
- Hank's unemployed family, who lives at his house. There's Hank's mother, who constantly attacks Hank for loving his job and not being a people-pleaser. There's Lillian Rearden, Hank's wife, who at least criticizes Hank more cheerfully than his mother does, and dutifully provides her body for his sexual needs. Finally there's Philip, who has "always placed the public good above any personal consideration" and contributes his time and Hank's money to various organizations including "Friends of Global Progress" (which Lillian also supports) in their crusade for the "Equalization of Opportunity Bill". This bill, to Hank's horror, would forbid any person or corporation to own more than one "business concern", but Hank can't believe such a bill would ever pass (you and me both, Hank!)
- The incompetent and vaguely socialist businessmen. Most notable is Orren Boyle, president of Associated Steel, Rearden's competitor, who is well-connected in Washington. These businessmen wouldn't dream of saying they are in business to make money; in fact, some businessmen in the story repeatedly tout that they have never earned a profit, and fall over each other to show how not-greedy they are. The associated companies tend to have collectivist-sounding names like "Amalgamated Switch and Signal" and "United Locomotive Works". How they survive without profits isn't made clear, but probably has something to do with Washington, DC.
- The honest, competent and productive businessmen: Ellis Wyatt of Wyatt Oil, Ken Danagger of Danagger Coal, Dwight Sanders of Sanders Aircraft, Andrew Stockton of Stockton Foundry, Dan Conway of Phoenix-Durango, Ted Nielsen of Nielsen Motors, Lawrence Hammond of Hammond Cars, Roger Marsh of Marsh Electric. Most of them are in Colorado, which has the fewest laws, and they begin to vanish, one by one, usually after decisions in Washington cause devastating shocks to their business.
- Wesley Mouch, who by the middle of the book seems to have more power than the president of the United States would have, if the president existed. The book never mentions a president, only a "Head of the State" (a seemingly minor character known only as Mr. Thompson, whose first brief appearance is halfway through the book). Rand explains in Chapter 2 that "nobody ever paid any attention to Wesley Mouch," but then Hank Rearden, who knows nothing about Washington except that all the big businesses have a "Washington man", hired Mouch as a lobbyist. Soon afterward, Mouch is the "Top

Co-ordinator of the Bureau of Economic Planning and National Resources". The story of Mouch and his implausible rise to power is finally explained in one paragraph in Part 2 Chapter 6.

- The elite Washington-friendly, irrational/anti-business elites, whom we first encounter when Lillian Rearden invites them to her anniversary party: Bertram Scudder, Dr. Simon Pritchett and Ralph Eubank. You can tell they are "intellectuals" because everything they say is ridiculous.
- Dr. Robert Stadler and Dr. Floyd Ferris of the State Science Institute. Robert Stadler, the official head of the SSI, was "the greatest physicist of his time" at the age of 30 and once said "Free scientific inquiry? The first adjective is redundant". At age 40 he endorsed the establishment of a State Science Institute, pleading "set science free of the rule of the dollar." Aside from this, Stadler is a figurehead who doesn't care about politics or life outside the realm of science. Meanwhile, Floyd Ferris, "co-coordinator", is a rotten pseudo-intellectual who seems to be the true head of the SSI.

As you can begin to guess, Rand had essentially two kinds of character: those of positive affect, and those of negative affect. Almost everyone is either a Rational Man of Action Who Seeks A Life of Joy (usually book-smart), or an Irrational Buck-Passer Who Blames Others For His Problems While Mumbling About Fairness And Morality (occasionally book-smart). In Part 1 of the book, the good guys rarely speak of morality while the bad guys speak of it frequently. Eventually the good guys develop a moral code based simply on thinking through their own intuitions (which are informed by their experience and their distaste for the bad guys). The evidence vindicating their philosophy is simply the obviousness of how wrong the Buck-Passers are, and the obviousness of how right the Men of Action are. (Buck-passing is so common that phrases similar to "I couldn't help it" appear 38 times, "blame" appears 75 times, "fault" 29 times.)

Note that Lillian, James, Ferris, Stadler and Hank's mother are quite different from each other in terms of their speech and skills, but they are all mostly devoid of likeable qualities. Good characters like Dagny, Rearden, Wyatt and Cheryl Brooks likewise vary, but end up with identical values. Both types are selfish, one openly, one secretly. Both types think of themselves as victims of the other type. No characters have any meaningful altruism; those who claim to possess any are just lying to themselves.

On the whole, Rand's "bad" characters are something like Joe McCarthy's [outgroup](#) at the bottom of his [death spiral of hate](#).

No black or asian people are mentioned in Atlas Shrugged, and there are few women of significance aside from Dagny and Lillian. One might hope that Rand or her future world is "color-blind", but I don't think so: at one point Boyle casually uses the word "Spic" when speaking to James.

Review

I have to credit Rand for having *some* understanding of evil and irrationality. Rand at least understood that every man is the hero of his own story - her villains don't believe they are

villains, and would (almost) never stoop so low as to use physical violence to get what they want. This, even though Rand clearly and frequently labeled everything about them as “evil”, and every government policy as the use of “force” - if you’re familiar with American libertarianism, you can’t fail to notice how close Rand’s worldview is to theirs. Indeed, the central plot clearly demonstrates, albeit with few details and low plausibility, just how forceful and destructive government policy can hypothetically be.

Rand also wrote skillfully the foolish, exaggerated doubletalk of her villains. I was genuinely impressed, because I have no idea how to speak like the most mentally stunted rabble on YouTube, while Rand seems to have done it effortlessly — though of course it is only the Buck Passers who ever speak so illogically.

However, as I read her book, I felt like one of her heroes, who at first is dimly aware of something wrong in the world, something that makes them uneasy, something that can’t possibly be true because it is such an unspeakable evil — but which, in the end, turns out to be completely real and every bit as bad as they could have imagined. In the same way, I became aware of something deeply disturbing in her writing, dimly at first, but something I could only grasp fully around the three-quarter mark.

In the novel, the great evil is something that superficially resembles communism (I guess she would call it “collectivism”; at one point she even uses the word “irrationalism”). Without explanation, this evil has conquered the government, the public school system, academia, and the media. But the great evil I saw was something very different, very dangerous, and very... well, very common, ordinary, and banal.

You see, at first I thought that Rand’s story was about some kind of conspiracy, a giant power play. I thought that she was hiding some great secret that would be revealed later in the book, something that would make sense in a world that I could imagine.

By the end of Part 2, I recognized the book’s great flaw as one big logical fallacy. If I could have described the book in two words, they would have been “false dichotomy” — or perhaps “strawman argument” — because there *really are* only two kinds of characters, Rand’s ideal human, and its opposite. Since no other kind of person exists in her world, there is no Plato for her Aristotle: no one rational enough to argue against her objectivist ideology or poke any holes in her arguments.

TV villains routinely do things that don’t make any sense: the villain who repeatedly murders his own lieutenants, for no other reason than to demonstrate evilness; the villain who kills someone for knowing about the existence of incriminating documents, instead of simply burning the incriminating documents.

Rand’s villains follow the same pattern, but with senseless goals instead of senseless violence, and so the government issues “directives” that would make no sense to most people: they have been chosen for their evilness. And not just any evilness, but specifically those Rand opposes, things like redistribution of wealth, irrationality, red tape/regulations, and “fairness”. Red tape and regulations are often framed as reduced freedom, so it’s notable that Rand never describes them that way (she simply says “red tape” and “regulations”), but I don’t know what significance to attach to a *lack* of framing. And Rand tends to distort “fairness” (and “equality”) into something superficial, absolute, uneven, and *unfair*, ensuring that it is unworkable, demotivating, and hurts some people and companies dramatically while benefiting others. The

villains (except Dr. Stadler) can figure out how to scheme, connive and control other people, but their brains fail at any other task. In particular, none of them can fathom Rand's ideology.

The directives are also arbitrary, undemocratic, foolish, destructive and unfair, but these are not the kind of adjectives Rand uses explicitly or spends the book ranting about; instead, such negative attributes play an implicit role, bestowing negative affect on the Buck-Passers.

Consider these directives issued at the end of Chapter 10:

The railroads of the country were ordered to reduce the maximum speed of all trains to sixty miles per hour — to reduce the maximum length of all trains to sixty cars — and to run the same number of trains in every state of a zone composed of five neighboring states, the country being divided into such zones for the purpose. (*red tape, charity toward labor unions, superficial fairness*)

The steel mills of the country were ordered to limit the maximum production of any metal alloy to an amount equal to the production of other metal alloys by other mills placed in the same classification of plant capacity — and to supply a fair share of any metal alloy to all consumers who might desire to obtain it. (*red tape, superficial fairness*)

All the manufacturing establishments of the country, of any size and nature, were forbidden to move from their present locations, except when granted a special permission to do so by the Bureau of Economic Planning and National Resources. (*red tape / regulations, irrationality*)

To compensate the railroads of the country for the extra costs involved and “to cushion the process of readjustment,” a moratorium on payments of interest and principal on all railroad bonds — secured and unsecured, convertible and non-convertible — was declared for a period of five years. (*redistribution of wealth*)

To provide the funds for the personnel to enforce these directives, a special tax was imposed on the state of Colorado, “as the state best able to assist the needier states to bear the brunt of the national emergency, “ such tax to consist of five per cent of the gross sales of Colorado's industrial concerns. (*redistribution of wealth*)

When I read this, I thought: give me a break! The U.S. is having economic difficulties, so the solution is to **cut** production? To shift trains from areas of high demand to areas of low demand? To arbitrarily punish railroad bondholders?

It was... just... too... stupid. Could she think that the Buck-Passers really would push through a plan this idiotic, with no concern for the next election? The book later indicates that the directives were the result of trading favors and power plays in Washington, but there is no explanation of how *special emergency powers* would be used only to grant a grab-bag of demands by special interest groups that will obviously make the emergency worse. Does Rand expect her *readers* to believe something this stupid? Well, the book was influential, so I guess she did, and I guess she was right. And I guess this is forgivable, given the idea Rand had in her head that the entire United States was pervaded by some sort of naïve Marxism. (She communicated this only subtly at first, which confused me, but perhaps right-wing audiences could pick up her signaling immediately.)

By the time I reached Part 3 Chapter 4, it finally dawned on me: she was not hiding any great conspiracy, and there would be nothing to finally explain the behavior of her characters sensibly. Consider this snippet from Part 2 Chapter 1:

Then, Wesley Mouch had issued another directive, which ruled that people could get their bonds “defrozed” upon a plea of “essential need”: the government would purchase the bonds, if it found the proof of the need satisfactory. There were three questions that no one answered or asked: “What constituted proof?” “What constituted need?” “Essential— to whom?”

One was not supposed to speak about the men who, having been refused, sold their bonds for one-third of the value to other men who possessed needs which, miraculously, made thirty-three frozen cents melt into a whole dollar. [...] Then it became bad manners to discuss why one man received the grant defreezing his money, while another had been refused.

Rand sometimes uses the kind of language you see in partisan editorials. How many editorials have you seen saying “no one is asking” a question that the editorial itself, and many others like it, are asking? “One was not supposed to speak about”? “it became bad manners”? Rand is attacking political correctness 40 years before it became fashionable!

But it’s more than that: she uses statements like these *as statements of fact*. This is not a mere smear piece; it’s a novel to promote her worldview. She never tried to smear communism or “collectivism” — she won’t even use those words. Instead, this is *literally* the way she viewed the world.

She wrote a book the length of *nine ordinary novels*. Can you imagine? Can you imagine how much of her life she much have sunk into this? To write something so empty, with villains so completely ridiculous, in a world where the people in power *consistently* do things don’t make logical sense? To fully invest in the premise that communism/collectivism is the literal opposite of rationality?

The scary realization I came to, finally, was how vacuous the book really was. Not her heroes, mind you — her heroes embodied the philosophy of her life, a life that she lived earnestly and honestly. But her villains! Holy crap! Eventually, she fleshed out her villains in great detail — never to the same extent as her heroes, but still, the level of detail leads me to believe that she *truly believed* that people like this are commonplace.

Oh, maybe she understood that in real life there are no *perfectly* horrible people. But I think she thought that that there is *only one* perfect good and perfect evil, and that her novel describes them both. She saw human beings, not as the complex nuanced variety that they are, but on a simple continuum from “good to evil”, “objectivist to collectivist”, “capitalist to communist” and “rational to irrational” respectively. To the extent that her characters have any secondary dimensions that can vary independently of the first, those dimensions are intellect, an ability to “talk straight” (heroes always talk straight, but certain government-aligned characters can also do so: Dr. Stadler, Fred Kinnan and maybe Mr. Thompson), and sexual liberation (only Hank Rearden was simultaneously good and not sexually liberated, but this changed as the book went on).

Given the extreme value she placed on honesty (even many bad characters are relatively honest), I think she was describing her *actual* belief about how evil socialists *actually* are, and I think the world in her book was what she feared would *actually* happen if they *actually* controlled academia. The book drops many hints, never in any detail, that there have been generations of children indoctrinated by schools and colleges, in essence, *not to think*.

Furthermore, given that her future world contains no technological advancements, her premise might have been that they were *already* in control of academia by 1957 (however, the story is incoherent in this regard: five amazing — and probably impossible — new inventions appear over the 12-year course of the story, which I cannot reconcile with the stagnation that came before). Google's first result for "what Rand believed about the real world" backs this up:

Rand herself stated, "My characters are never symbols, they are merely men in sharper focus than the audience can see with unaided sight. ... My characters are persons in whom certain human attributes are focused more sharply and consistently than in average human beings"

While Rand's heroes are certainly self-centered and somewhat greedy, they often seem less greedy and more fair than her villains, who lust after power, which is not described as "greedy". Rand tries to convince readers that greed is good by comparing it to something worse that is not *called* greed; similarly, she labels foolishness as fairness so that fairness looks bad.

Rand's writing is common, ordinary, and banal in the sense that it often sounds like I would expect an anti-communist editorial from the 1950s to sound. What does the book have to say about the teachings corrupting the men of her future world? It's the same pithy phrases those editorials would have used to describe the teachings of communism (and irrationalism, if that were a real thing), things like "individual ability is of no consequence" and "the law of causality [is] a superstition" — that's all, *nothing* else. It's as if everything she knows about her enemies came from short "opinion" pieces in local conservative newspapers.

The great flaw of the book, then, is not Rand's ideology per se, but her approach to reasoning. She created the weakest strawmen she could think of, with the weakest backstories — like houses built on stilts — and then acted as if knocking them down was a victory for her rationality. Among all her reasoning mistakes, this was the cornerstone: a mistake that blocked her from noticing any other mistake. But I guess I can forgive her, as 1957 was more of an epistemic wild west. As Yudkowsky [put it](#),

Max Gluckman once said: "A science is any discipline in which the fool of this generation can go beyond the point reached by the genius of the last generation." Science moves forward by slaying its heroes, as Newton fell to Einstein. [...] An aspiring rationalist in 2007 starts with a huge advantage over an aspiring rationalist in 1957. It's how we know that progress has occurred.

Blind Spots

One other systematic error jumps out at me: Atlas Shrugged is written as though key elements of human society are invisible to Ayn Rand.

The book is overwhelmingly dominated by individuals - individualists on the one hand (good, thing-oriented), and collectivists on the other (bad, people-oriented). The interesting part is that while she considers *physical* systems in some detail, *social* systems and institutions are thoroughly ignored.

She mentions a few times the churn of trade: how ore is moved by rail to steel mills, how ball bearings are made and shipped to factories that use them, how trains and rails and sidings and signals and platforms help ship goods around the country. Rand seems to understand both the importance of heavy industry and its functional details, and she does a better job than I could of casually explaining them.

Yet political structures seem entirely off her radar: the electoral college, first-past-the-post elections, the court system, the president, the constitution, the bicameral congress, city councils, the ways candidates can and cannot get campaign funding - such structural elements and institutions of government are never discussed, as though Rand believes they simply don't matter. For example, the only purpose of the national "legislature" is to hand over emergency powers to Wesley Mouch and the Head of State, with no explanation of how or why this happens.

The story seems to be meant as a cautionary tale, so it is strange that the book never more than hints at any altered political structure in the United States, and never has anything to say about what happened to cause the changes. It's all "causeless" (to borrow a word she used 29 times in the book).

Nongovernmental institutions are largely ignored as well. Not a single news organization has a name; they are simply "the press" and they seem uniformly to speak as one voice with the Men in Washington about "fairness" and "greedy corporations" and "helping the needy". There is a "union" of railroad companies, but their main role in the book is to vote against competition with an "anti-Dog-eat-dog rule" (no motive for passing this rule is given at first, but eventually there is a vague explanation that it was a backroom deal among the Buck-Passers). The book's main role for colleges and universities is for them to close as the economy collapses. Banks don't really play a role either. Nonprofit organizations are rare, and useless at best. Workers' unions play a minor role supporting foolish, but ironically self-interested policies.

Even the ancient institution of the *family* is ignored. Slowly, virtually every competent leader of every major company "vanishes" (quits and hides) as well as numerous important managers, and later, most industry veterans break the law when they quit and hide. *What about their families? Why don't they have families?* The story wouldn't work very well if they did. (Exactly one mother who "vanished" is mentioned in the book; she says her "profession" is motherhood, but there is no mention of who pays her salary. She's married, so we are left to wonder how this "profession" would work for single moms.) Ditto for friends; the story relies on the "good" characters having such extreme individualism that they will gladly (and I mean *gladly*) leave friends and coworkers behind.

It is not until Part 2 Chapter 6 that we learn that Rearden Steel actually has a union, and this is mentioned only because the head of the union quits:

The first man to quit at Rearden Steel was Tom Colby, rolling millforeman, head of the Rearden Steel Workers Union. For ten years, he had heard himself denounced throughout the country, because his was a "company union" and because he had

never engaged in a violent conflict with the management. This was true: no conflict had ever been necessary; Rearden paid a higher wage scale than any union scale in the country, for which he demanded—and got—the best labor force to be found anywhere.

Given the decline of unions in the decades since 1957, it's interesting that Rand takes unions so much for granted that the existence of a union at Rearden's company isn't important enough to mention for several hundred pages. Notice also that Rand's hero pays high wages. No doubt this is not because Rearden is "fair" (one of Rand's four-letter words, alongside "practical") but because he wants to buy "the best" people. Yet wouldn't he be able to pay lower average wages if there wasn't a union, and keep more profit for himself?

Hank Rearden publicly tells everyone that he only cares about making money, and does not shy away from lavish spending on himself and family or friends, but it's clear that Rand's heroes are not nearly as greedy as they could have been. Atlas Shrugged has nothing like a telecommunications firm using a monopoly on telephone wires to keep rates high, nothing like an Apple corporation extracting 30% of sales on all apps made by other companies, nothing like a John Deere corporation trying to prevent "unauthorized" parties from repairing tractors, no patent trolls, no old-money people living well off interest from a diversified portfolio without working for a living.

Protestors Criticized For Looting Businesses Without Forming Private Equity Firm
First - [The Onion](#)

This is not the kind of book that would consider whether it is a grave injustice for the government to use tax money to build paved roads in front of the houses of the taxpayers, allowing any yahoo with a car to freely drive past houses that don't belong to him, on roads that don't belong to him, or whether the government can ever be justified in putting sewers or fresh-water pipes or copper wires or fiber-optic lines under these roads using (gasp!) funds from income taxes.

Nor does the book consider whether children have a "right" to an education and, if so, whether it's okay for the government to pay for educating the children of the unemployed. Should there be standards for food and drug safety? Rand never mentions the FDA or anything like it. Should there be a government-assisted health care system? No one ever gets a serious illness (or mental illness) in the book, let alone someone who just lost a job. Would the private sector ever pay for a national census, or an [LHC](#), or research on ancient languages (which might offer insights into [human theory-of-mind](#))? What about problems like overfishing the Atlantic ocean, or global warming?

The pirate Ragnar Danneskjöld, one of Rand's "good" characters, treats income tax as simple theft, and his friends tacitly agree. We can speculate that he would insist that the water pipes must be built and owned by one person or corporation, to be rented out to those who depend on it at any price the market will bear; that children should only be educated if the necessary funds are given freely by private citizens; that the street is a proper home for a schizophrenic who loses a job; that if people want food safety standards or don't want overfishing, they should *somehow* (how?) come up with a contract among themselves to arrange it. But Rand's *actual* solution to such questions is to ignore them.

Instead, Rand sticks to the easy question: which is better, a “greedy” unionized steel manufacturer trying to sell the country cheaper, stronger metal than everyone else, or filthy collectivist scumbags making secret backroom deals in an effort to steal from the rich and give to the lazy? Ironically, the structure of Rand’s book has exactly the effect she decries in her enemies: it shields her from having to think about any hard questions whatsoever. She describes villains whose minds are twisted into pretzels by their illogical ideology, but her book’s entire world is twisted into a pretzel to fit *her* ideology.

The story provides glimpses of people who are able and willing to work hard, but end up poor and destitute anyway, but it never asks itself whether Rand’s preferred form of government would sometimes produce such outcomes and whether, if so, it would be just. Moreover, the story relies on the secrecy and perfect conviction of the deserters; it never tries to figure out whether her utopic ideas would work in any other context. I won’t spoil the ending, but the book also never considers whether heroes acting morally could have engineered a *vastly* better outcome than the one in the book. A dramatic ending is valuable in fiction, but her heroes *never consider* other courses of action; I think Rand was [motivated not to look for other options](#).

I appreciate Ayn Rand’s understanding of people who care not about trading favors, about smalltalk and social pleasantries, about beating others down, people whose dream is to do good work, to accomplish great things, and to live happily — people like me.

Yet it strikes me that at least two of the heroes of the story, Francisco d’Aconia and Dagny Taggart, are people who inherited a fortune and a grand opportunity to succeed in life. And we see the same phenomenon in real life, of course, in people like the Koch Brothers or Donald Trump or [these 16 billionaires who inherited their money](#). Could Francisco and Dagny have been as important as they were, had they not been born into a family dynasty? In the world of Atlas Shrugged, I don’t think so, but this is a question that Rand’s heroes never ask. What about a hypothetical objectivist utopia? At least among those who won an equal amount in the genetic lottery of life, does Rand imagine there could be anything close to equality of opportunity in her ideal world? It’s hard to tell.

Probably the hardest question the story asks of its readers is whether adultery can be condoned in a loveless marriage. Rand’s answer is a resounding “yes”. Rand would probably think it better for the would-be adulterer to inform the other spouse of his intended action in advance, but it’s not directly stated. Rearden keeps his affair secret from his wife and everyone else, but he only handles it that way because he originally believes in a 1950s-style moral code that condemns sex and is divorce-averse, viewing marriage vows as a solemn, joyless duty that one cannot be seen breaking. Rearden eventually comes to his senses and realizes that his moral code had been wrong, and that his affair is right and just after all.

Closing remarks

Eventually I decided that the existence of such a strange book was not surprising after all, nor was its popularity. The world contains many unusual individuals. Ayn Rand was one; so are many of the people reading this. But whether they become famous and influential depends on

whether their message resonates with people, and Rand combined the respectable appearance of a philosopher with beliefs that were similar to popular ones.

“There is always an easy solution to every human problem—neat, plausible, and wrong.” - Misquote of H. L. Mencken

Successful scientists and philosophers have been attracted to finding simple explanations and rules, and there are good reasons for this and good reasons to seek them out. Ayn Rand found an attractively simple philosophy of savages versus heroic men, which rubbed off onto American libertarianism.

“Collectivism is the tribal premise of primordial savages who, unable to conceive of individual rights, believed that the tribe is a supreme, omnipotent ruler, that it owns the lives of its members and may sacrifice them whenever it pleases.” - Rand’s *Capitalism: The Unknown Ideal*

“My philosophy, in essence, is the concept of man as a heroic being, with his own happiness as the moral purpose of his life, with productive achievement as his noblest activity, and reason as his only absolute.” - Appendix to *Atlas Shrugged*

Scott Alexander taught me another, even simpler philosophy that I think works much better: consequentialism, and its more specific incarnation, [utilitarianism](#). A [FAQ that explains consequentialism](#) was the first piece of Scott’s writing that I ever saw; without it, I might not read SSC today (er... what’s that? it’s “ACX” now? And the X is pronounced “ten”? Okay, whatever.)

Like the [Mandelbrot set](#), consequentialism is a simple idea with infinitely intricate, well, consequences. For example, I have learned the value of [deontological ethics](#) through the lens of utilitarianism, in light of the practical capabilities of human beings and the impracticality of enumerating every possible action and predicting the consequences of each.

But, hell, this is just a book review, so I won’t get into it. I will say, though, that if you’re sympathetic to objectivist and libertarian philosophies, you’d better give consequentialism serious consideration as well. And for those who have *already* made up their mind, just make sure that this quote does **not** sound like you:

“I have held the same philosophy I now hold, for as far back as I can remember.” - Afterword, *Atlas Shrugged*

This might have something to do with her father’s pharmacy being [seized by Bolsheviks](#) when she was 12. But regardless of what your beliefs are, *always having held the same beliefs* is deeply suspicious: we are not born with a map of reality in our minds, and the first map we build is sure to contain errors. If you’re still stuck on the first interpretation of morality you ever learned, make sure that you at least know [How To Actually Change Your Mind](#).

BadAtlas Shrugged by Ayn Rand

[SLP]

Spoilers ahead; content warning: mention of suicide

I.

I read *Atlas Shrugged* because my library closed down for coronavirus, and at the little free library at the church near my house, the choices were this or Nancy Drew books. After spending several weeks reading it, I have come away from Ayn Rand's magnum opus with the sense that no one else has ever actually read it before.

I made my roommates nervous by cackling while reading it, but the truth is that Rand is much funnier and hornier than I could have ever dreamed. The main bones she picks are actually pretty reasonable, and the book ends up being far more critical of capitalism than I expected. In particular, she spends a lot of time making extremely insightful, defensible critiques of government corruption. In this regard, Rand is [one cookbook](#) away from Elizabeth Warren.

While I understand the appeal of this novel, Rand's values are ultimately inconsistent and surprisingly naïve. Even reading generously, the failings of her objectivist ideology are pretty evident. Still, since I've already sunk the time into reading 1000+ pages, I hope to help you have a more interesting, nuanced point of view when you pretend to have read it at a party in the future.

II.

Who is John Galt?

This, according to my 50th Anniversary Edition of *Atlas Shrugged*, is the "immortal query" that Ayn Rand created as the "perfect artistic form to express her vision of existence." In the novel, this rhetorical question pretty much means "who knows" or "why bother asking questions when you'll never find out the answer."

I finished *Atlas Shrugged* with many questions, but this one I can definitively answer. John Galt is our titular character--the Atlas who held the world on his shoulders, and whose shrug sent the world crashing down around him. A physicist and philosopher of prodigious intellect, Galt became frustrated by the incompetence of the world and decided that he would bring the system to its knees by organizing a strike of all the worthy people. These are mostly industrialists but also include a banker, a judge, a composer, a writer, and a movie star. His stated goal is to "stop the motor of the world," and he will know this has been achieved when the lights of New York go out.

While Galt is the hero, our protagonist is Dagny Taggart, Vice President in Charge of Operations for Taggart Transcontinental, who splits most of her time between saving the railroad from falling apart in the hands of her brother James and *schtupping* steel magnate Hank Rearden. She also sometimes stands alone at dinner parties, falls asleep under a portrait of her grandfather and Taggart Transcontinental founder Nathaniel Taggart, or feels conflicted about her childhood friend and former lover Francisco D'Anconia, the heir to a copper fortune who has now turned into an international playboy. Oh, and she's hot--even (or perhaps especially) when wearing a suit.

Guided by an extremely judgemental narrator (Rand's insults are truly unparalleled), the book takes the reader on a wild ride to see a giant calendar in the sky, a plane chase, a love triangle, a secret government science project, a marriage between a railroad president and a shopgirl, a 40-page long monologue, a lab that implodes if you force your way in without a key, a daring rescue, and a helicopter ride above New York City as its lights flicker out.

The primary story arc is not fully revealed to the reader until Part III of the book, when Dagny crash lands in the remote Colorado valley. Galt has chosen this valley as the place for his "strikers" to hide in, and they've chosen to call it *cue groans* Galt's Gulch. Here we finally meet John Galt, we and discover that all of those characters Dagny watched disappear, along with some others she had only ever heard of, are here together, building their own community away from the "looters and moochers" who want to take it all away from them.

(The reason she ends up crashing in the first place is that she's trying to find a young physicist she hired to try to reconstruct a revolutionary motor that harnesses static electricity, which is now used to power the strikers' whole settlement. This is Galt's great invention, and is one of a few interesting bits of sci-fi in the novel, including a ray-shield that makes Galt's Gulch invisible to outsiders and the government's Project X, which creates a weapon of mass destruction that uses ultrasonic sound. I assume this last one is loosely based on the Manhattan Project.)

It's actually already pretty obvious what's going way before you get to the Gulch, but Rand still spends quite a few words explicitly revealing everything and having Dagny be appropriately surprised. The only really surprising thing to happen here is that Dagny's love triangle with Hank Rearden and Francisco D'Anconia resolves because she falls for John Galt instead, who has beautiful hair that shines like gold and/or copper depending on the passage.

(Have I mentioned that Rand really, really loves Gold? If you didn't already know, "The code of competence is the only system of morality that's on a gold standard." The banker in the town is named Midas Mulligan and they have their own currency--gold coins! They make the best cigarettes ever and put a little gold stamp on them! There's a giant dollar sign made entirely of gold!)

Eventually, the world outside crumbles. Dagny is with John Galt and Ragnar and his hot ex-movie star wife and some others are all together planning their future return to the world they abandoned and how they won't be nice to each other in business deals. As the wind blows through his luscious locks, [John Galt traces a dollar sign in the air](#). Fin.

When I started reading, I thought this book was a philosophical and economic treatise beloved by conservatives, and so it was kind of weird to discover that it's a book about... creating an intentional community? Apparently no one takes this seriously (okay, fine, a few people do, but it seems to [fail spectacularly](#)), and her fans like Paul Ryan and Mark Cuban are still out here participating in society. Galt's Gulch doesn't hold any water. Still, it remains incredibly popular and incredibly divisive.

What gives?

III.

Part of the answer is that *Atlas Shrugged* is way more readable than any other primary source philosophy text I've read. The book is longer than it needs to be (way, waaayyy longer), but Rand's prose is accessible and none of her main points rely on obscure terminology. Plus, she repeats herself so many times that even if you skim you'll still pick up all of her main points.

If you were wondering why I kept mentioning how hot they all are, it's because Rand keeps on mentioning how hot they are. The schtick of this book is that none of the characters are remotely realistic. There are only heroes and villains--people are either strong-willed geniuses with perfect muscle tone and glistening hair or incompetents with hunched shoulders and weak mouths.

Atlas Shrugged is also much [steamier](#) than the traditional ivory-tower philosophizing. The sex scenes aren't particularly pornographic, but everyone knows that [sex is cool](#) and the libinidal energy Rand brings holds the story together through some weak moments. Since everyone is a hero/villain from the get-go, there is not much character development or engaging interpersonal drama besides Dagny's relationships with Hank/Francisco/John Galt, but there's enough intrigue and drama in the romance plotline to string the reader along when the main plot lags. I primarily associate *Atlas Shrugged* with young men, and it makes sense that this demographic would be particularly interested in an ideology that involves getting laid.

(Unfortunately, there isn't any crazy sex on a speeding train--if this is your cup of tea, I would suggest you instead watch the [Trotsky mini-series](#) which is currently available on Netflix and begins with an absolutely bonkers sex scene on a train speeding through the Russian wilderness.)

Rand's prose really shines when she's being mean. Let's examine a few choice insults:

- "Men with gelatin eyes, rubber voices, spiral-shaped convictions, non-committal souls and non-committing hands"
- "The face of his attorney, an elderly man of the old-fashioned school, wore an expression that made it look like he longed to take a bath"
- "Schoolroom voice and bar room mouth"
- "Her voice sounded as if it were falling in drops, not of water, but of mayonnaise"

My key takeaway from reading *Atlas Shrugged* in 2020 is that it's a real shame we never got to see the kind of takes Rand would drop on Twitter. She would have had some choice things to say about those pictures of Mark Zuckerberg sucking up to Congress in a too-big suit.

Ayn Rand also understands what it's like to put up with bullshit; watch pretty much any of her interviews and the frustration is palpable. Being competent and dealing with incompetent buffons sucks, and it's cathartic to read about a group of people (who are morally righteous and get laid) who stuck it to the powers that be and brought it all crashing down. *Atlas Shrugged* is revenge fantasy in the vein of *The Count of Monte Cristo* or *Gone Girl* or *Kill Bill*. Aspects that at

first blush make the story unrealistic--black and white morality, larger-than-life characters--ultimately contribute to an epic feel that allows it to transcend everyday annoyance to seem exciting and transformative.

Hidden within the revenge fantasy are some interesting kernels of self-help. Dagny's relationship with Hank Rearden is transformative and compelling not only because she says hot things to him like "You'll have me any time you wish, anywhere, on any terms" but because she is lonely and misunderstood and with him has found the importance of mutually supportive relationships. Especially when we are given the contrast between Hank's really shitty relationship with his wife Lillian, the importance of good company becomes unmistakable.

Camaraderie and companionship make things ok, even when the looters are stealing your intellectual property and passing regulations that makes it practically impossible for you to run your transcontinental railroad. For all his ability as a physicist and inventor, the thing that gives Galt his power is his ability to find and convince like-minded people to join his cause. If you feel like a misfit, it's because you're around the wrong people. But don't worry--they're out there, even if they're inaccessible to you because they're all hiding out somewhere in Colorado.

The best part about Rand's philosophy is her insistence that each individual think for themselves and take responsibility for their actions. John Galt speaks of three fundamental values: reason, purpose, and self-esteem. Objectivism gives each individual the authority to determine their own values and worth. Doing things that make you happy is virtuous. The advice to take ownership of your own life is pretty universally recognized as good, and *Atlas Shrugged* makes a compelling case for the importance of believing in yourself and finding meaning in taking responsibility.

This, I think, is the really beautiful thing about Rand's philosophy that explains a significant portion of her following. Those other people who are trying to tell you how to live your life, who are making you feel ashamed about the things that bring you joy? Fuck 'em. At its core, the story of the strikers is about helping you, dear reader, realize that not only do you deserve nice things, you even get to decide what things you think are nice.

IV.

Hence the famous Randian praise of selfishness.

And it is true that Ayn Rand tells us to be selfish. She declares that every man “is an end in himself, not a means to the ends of others; he must live for his own sake, neither sacrificing himself to others nor sacrificing others to himself; he must work for his rational self-interest, with the achievement of his own happiness as the highest moral purpose of his life.” This seems pretty straightforward, until we start to look at how the rest of the characters are acting.

The heroes are certainly selfish, but so are the villains. Dagny’s brother Jim, Hank Rearden’s former lobbyist Wesley Mouch, and a whole slew of Washington looters and phonies with silly names like Kip Chalmers and Balph Eubank are all out there looking out for themselves and their bottom lines. Jim uses his connections to advance legislation that crushes his railroad competition in Colorado. Wesley Mouch betrays Hank Rearden to gain influence in Washington. Sure, they talk a big talk about the “public good,” but at the end of the day, everyone is lining their own pockets.

The critique she’s offering is not one of altruism or even socialism. She does discuss foreign “People’s States”, but there are no apparent welfare systems or social services. They’re not even doing this to try to distribute, but rather to double down on industry influence to use the government as a tool to maintain existing systems of inequality.

Let’s take a look at some of the government actions in the novel Rand finds so abhorrent. Directive 10-289 attempts to stave off economic decline by banning change. It is explicitly written to create economic stagnation to benefit existing monopolies. Workers aren’t supposed to change jobs, businesses aren’t supposed to close, nothing new should be invented.

When the Americans lobby their South American friends to nationalize Francisco’s family’s copper industry, they do so because they personally stand to benefit enormously. Even the nationalization of the steel industry, which they do *by using government agents to infiltrate the steel mills and instigate riots so the state can step in to protect the safety of the mill owners*, occurs in order to save the business of Orren Boyle, whose Associated Steel cannot compete with Hank Rearden but has strong enough government influence to save his business with that.

The Anti-Greed Act, Equalization of Opportunity Act, and--my personal favorite--the Anti Dog-Eat-Dog Act are all anti-competition, but with the goal of protecting the existing systems and entrenched capitalists. These aren’t acts of redistribution, but are the people already in power consolidating their wealth. Despite her reputation, it’s clear that Rand isn’t just blindly pro-capital all the time. The Anti Dog-Eat-Dog Act isn’t even government action--it’s passed by the National

Alliance of Railroads because it benefits the more powerful members. These industrialists are choosing to self-regulate using their voluntary associations.

There's no alternative. All players are acting selfishly, and while they anti-free market, they aren't socialist. What Rand is truly concerned with the ways in which one individual's selfish pursuits interfere with another's ability to be selfish. This appears as corruption, which Rand charmingly calls "pull-peddling."

The most compelling bits of this novel come in the savage critique of corrupt politicians and trading favors. In Rand's world, supposed altruism and socialism are really covers for two things: blatant stupidity and justifying use of force.

The first real sin is always incompetence. Some things exist outside of our minds and our social circles. Corn and soybeans will rot if we don't have the trains to deliver them to consumers. The people Rand describes as looters and moochers don't have their shit together enough to do basic things like get trains to run on time or produce enough to meet all their orders or generally understand that getting things done. Instead, they are obsessed with relativism and fake news. As a result, they get in the way of everyone else who is trying to accomplish great industrial progress. What Rand really wants is a reorientation away from social games towards the external physical world, putting the objective in objectivism.

Rand is also particularly concerned with corruption in the justice system. Incompetents slide away from various truths that govern the world.

"What would he be able to prove? To whom? One could prove nothing to a tribunal that had no stated policy, no defined procedure, no rules of evidence, no binding principle--a tribunal, such as the Unification Board, that pronounced men guilty or innocent as they saw fit, with no standard of guilt or innocence"

Ultimately, this devolves into a post-truth system.

"When a court is not bound by any rules, it is not bound by any facts, and then a hearing is not an issue of justice, but an issue of men, and your fate depends not on what you have or have not done, but on whom you do or do not know"

You are who you know, and that determines what you can get away with.

In a chapter hilariously titled “The Moratorium on Brains,” Rand further explores how corruption and scapegoating within the government seeps into the workforce. Following Directive 10-289 and its creation of the Unification Board, employees need to be careful not to step out of line lest they be punished. This prevents independent thinking and promotes compliance above all else.

Politician Kip Chalmers is worried about his train arriving late to a rally, and so forces the railroad to use a coal-burning engine through a tunnel even though this will suffocate them. Dagny is busy repairing a path at a mountain retreat and is not available to save the day from feckless behavior.

While some workers know the danger, they shrink from their duty. They know that if they do not comply they will be punished. Their primary goal is to avoid blame and be allowed to continue to work to support their families. As a result, they willingly abdicate their ability to think for themselves. In this way, a corrupt system of justice leads to ever greater incompetence. Rand is not patently pro-capitalist and anti-worker. I think she honestly just doesn’t like most people at all, regardless of their station in life.

These concerns gain additional significance because the government is empowered to determine legal guilt. They create laws so strict that everyone can be made guilty forces individuals to engage in endless cycles of blackmail and influence peddling.

“Did you really think that we want those laws to be observed?...We want them broken. You’d better get it straight that it’s not a bunch of boy scouts you’re up against—then you’ll know that this is not the age for beautiful gestures. We’re after power and we mean it. You fellows were pikers, but we know the real trick, and you’d better get wise to it. There’s no way to rule innocent men. The only power any government has is the power to crack down on criminals. Well, when there aren’t enough criminals, one makes them. One declares so many things to be a crime that it becomes impossible for men to live without breaking laws. Who wants a nation of law-abiding citizens? What’s there in that for anyone? But just pass the kind of laws that can neither be observed nor enforced nor objectively interpreted—and you create a nation of lawbreakers—and then you cash in on guilt.”

Beyond the basic injustice, Rand is concerned that this system of government leads to more direct violence.

“But when a society establishes criminals-by-right and looters-by-law — men who use force to seize the wealth of *disarmed* victims — then money becomes its creators’ avenger. Such looters believe it safe to rob defenseless men, once they’ve passed a law to disarm them. But their loot becomes the magnet for other looters, who get it from them as they got it. Then the race goes, not to the ablest at production, but to those most ruthless at brutality. When force is the standard, the murderer wins over the pickpocket. And then that society vanishes, in a spread of ruins and slaughter.”

The villains use exploitative, forceful means to achieve their ends. This is the ultimate act of evil.

“Whatever may be open to disagreement, there is one act of evil that may not, the act that no man may commit against others and no man may sanction or forgive. So long as men desire to live together, no man may *initiate*—do you hear me? no man may *start*—the use of physical force against others.”

Clearly the use of force is the most dastardly thing imaginable. Except not really, because Rand’s heroes also use violence to achieve their goals.

Nathaniel Taggart is Dagny’s personal hero. She had a portrait of him in her office, she frequently stops to admire his statue in the Taggart Terminal. He represents the best of industry--the man who beat all the odds to create a transcontinental railroad and become absurdly wealthy in the process.

How did he do it? By resorting to violence on multiple occasions.

“It was said that Nat Taggart had staked his life on his railroad many times; but once, he staked more than his life. Desperate for funds, with the construction of his line suspended, he threw down three flights of stairs a distinguished gentleman who offered him a loan from the government.”

It gets better!

“It was said that in the wilderness of the Middle West, he murdered a state legislator who attempted to revoke a charter granted to him, to revoke it when his rail was laid halfway across the state; some legislators had planned to make a fortune on Taggart stock – by selling it short. Nat Taggart was indicted for the murder, but the charge could never be proved. He had no trouble with legislators from then on.”

Now I'm not sure if Middle West is a spin on midwest or if she means the middle within the west, but I am relatively sure this doesn't pass the self defense sniff test. We're not given a ton of context, but Rand certainly had no problem with being an asshole, and despite her pronouncements against violence she seems very willing to forgive it.

V.

The real problems arise because the looters are playing the game of the market while also simultaneously playing a game of political influence. Even if a new outsider can outcompete the old guard in the market, they will lose because people in charge of other avenues of power can use external forces to shut them down. But this isn't anti-capitalist either--this is a possible future of what capitalism looks like when it's allowed to slide from the city on the hill into the swamp.

Rand illustrates how the government has become the ultimate monopoly, and acts as a clearinghouse through which already established corporations can extend their influence. The lines between industry and government blur. Wesley Mouch, originally employed as a lobbyist for Hank Rearden, ultimately turns on him to pass the Equalization of Opportunity Bill, which forces Rearden to sell of many of his companies due to new restrictions on how many businesses an individual can own. In reward for doing so, he is given a cushy government appointment. Voluntary capitalist associations like the National Alliance of Railroads similarly become tools of oppression.

In the long term, capitalism consolidates power in a way that allows this corruption. There are no political parties in *Atlas Shrugged*. The bad faith actors aren't just politicians--they're also the industrialists. Instead of facing the consequences of the market and taking a hit from making bad bets, they turn to protective regulations and breaking kneecaps. Free markets are only free until tertiary power structures are erected to protect the winners of the first round.

Winners who accumulate capital and private property find that the best way to keep winning is to shut down the system that allowed them to rise. The people in charge are certain their investments will continue to accumulate value because they've rigged it to do so. A true free market is a danger to those in power because bets can be lost. This ends up sounding sadly prescient.

Newcomers simply cannot innovate their way out of the system. Ellis Wyatt pioneers a new process to extract oil, but is ultimately crushed by government regulations. Hank Rearden, who came from nothing to run the largest, most successful steel company, creates a new miracle metal known as Rearden Steel, only to have a law passed that forces him to share the intellectual property.

In a scenario where there is no John Galt leading a strike (i.e. real life), Rearden's only options would be to let his own business be crushed or to hop aboard the pull-peddling party. Before the strike begins to affect the economy in earnest, James is clearly outcompeting Dagny on the whole. Even if he's worse at running a railroad, he still wins because of his advantages using his influence with the Taggart Transcontinental board and politicians.

Rand does not take up how this system was put into place in the first place. Still, her story clearly illustrates how entwined economic and political oligopolies are created and perpetuated. One way is through familial inheritance--James Taggart is president of Taggart Transcontinental purely because he is a descendent of Nat Taggart. Even if all the spoils of the first round are justly distributed, the incompetent will always end up with the loot.

Rand clearly believed that biological family is not a good indicator of similarity. In several cases, a striker's abandoned business went to their brother/cousin/other relative who let it fall apart. Hank Rearden could not be more different from his brother Phillip, and boy does he hate his mother. Dagny and James are a textbook example of literary foils.

(Side note: I am driven completely mad by the fact that their parents essentially do not appear to the degree that they are never even named!!! Who are they!! What happened to them!!)

Rand thinks that Nat is perfect and James is evil, but fails to see that the same system that allowed Nat's wild success twists itself to prop up James and his looting a few generations later. What began as a competition in the market has become a competition in the swamp.

How do we prevent this deterioration? In what ways should we be deciding who gains influence during transitions of power? How do we ensure innovation can break through in the long run? How do we ensure the longevity of an enterprise? These are all interesting, important questions that I found myself asking while reading, and was ultimately disappointed that Rand doesn't address them at all.

The whole point of Atlas shrugging is that the only way to escape this system is to break it down so you can start over again. The system tends away from a free market. Rand does an amazing job of identifying the inevitable ills that come from uncontrolled political entropy and crony capitalism, but she fails to imagine a way to reboot the system without completely destroying it.

Of course, the dire consequences of incompetence are exaggerated in the universe. The economy only goes downhill as precipitously as it does because Galt and company are encouraging its fall. Friendly neighborhood pirate Ragnar keeps on sinking the ships carrying necessary shipments of copper and other materials.

Rand is surprisingly open to destroying property just to stick it to her enemies, though ultimately this reveals a key aspect in this world that allows the strikers to survive and succeed--boundless resources. The valley they choose for Galt's Gulch is ridiculously plentiful, offering excellent land for agriculture and a number of natural resources. Without this plenty, we simply don't have the option to start a standalone colony, or to destroy our world in order to cleanse it for a new beginning.

VI.

Boy does she destroy it.

Perhaps the most interesting question I asked myself while reading *Atlas Shrugged* was what to do with all those feeble-minded weak-chinned incompetents? The book ends 9 pages after the lights of New York go out, so we see very little of what the world looks like. But even before this, the prospects of most people were pretty grim. A lot of people presumably die, so let's take a look at a few of the people who definitely die in the aforementioned train crash. Rand gives us a nice long list of passengers and their various sins.

"It is said that catastrophes are a matter of pure chance, and there were those who would have said that the passengers of the Comet were not guilty or responsible for the thing that happened to them.

The man in Bedroom A, Car No. 1, was a professor of sociology who taught that individual ability is of no consequence, that individual effort is futile, that an individual conscience is a useless luxury, that there is no individual mind or character or

achievement, that everything is achieved collectively, and that it's masses that count, not men.

The man in Roomette 7, Car No. 2, was a journalist who wrote that it is proper and moral to use compulsion 'for a good cause' who believed that he had the right to unleash physical force upon others - to wreck lives, throttle ambitions, strangle desires, violate convictions, to imprison, to despoil, to murder - for the sake of whatever he chose to consider as his own idea of 'a good cause', which did not even have to be an idea, since he had never defined what he regarded as the good, but had merely stated that he went by 'a feeling' - a feeling unrestrained by any knowledge, since he considered emotion superior to knowledge and relied solely on his own 'good intentions' and on the power of a gun.

The woman in Roomette 10, Car No.3, was an elderly schoolteacher who had spent her life turning class after class of helpless children into miserable cowards, by teaching them that the will of the majority is the only standard of good and evil, and that a majority may do anything it pleases, that they must not assert their own personalities, but must do as others were doing.

The man in Drawing Room B, Car No. 4, was a newspaper publisher who believed that men are evil by nature and unfit for freedom, that their basic interests, if left unchecked, are to lie, to rob and murder one another - and, therefore, men must be ruled by means of lies, robbery and murder, which must be made the exclusive privilege of the rules, for the purpose of forcing men to work, teaching them to be moral and keeping them within the bounds of order and justice.

The man in Bedroom H, Car No. 5, was a businessman who had acquired his business, an ore mine, with the help of a government loan, under the Equalization of Opportunity Bill.

The man in Drawing Room A, Car No 6, was a financier who had made a fortune by buying 'frozen' railway bonds and getting his friends in Washington to 'defreeze' them.

The man in Seat 5, Car No.7, was a worker who believed that he had "a right" to a job, whether his employer wanted him or not.

The woman in Roomette 6, Car no. 8, was a lecturer who believed that, as a consumer, she had "a right" to transportation, whether the railroad people wished to provide it or not.

The man in Roomette 2, Car No. 9, was a professor of economics who advocated the abolition of private property, explaining that intelligence plays no part in industrial production, that man's mind is conditioned by material tools, that anybody can run a factory or a railroad and it's only a matter of seizing the machinery.

The woman in Bedroom D, Car No. 10, was a mother who had put her two children to sleep in the berth above her, carefully tucking them in, protecting them from drafts and jolts; a mother whose husband held a government job enforcing directives, which she defended by saying, 'I don't care, it's only the rich that they hurt. After all, I must think of my children.'

The man in Roomette 3, Car No. 11, was a sniveling little neurotic who wrote cheap little plays into which, as a social message, he inserted cowardly little obscenities to the effect that all businessmen were scoundrels.

The woman in Roomette 9, Car No. 12, was a housewife who believed that she had the right to elect politicians, of whom she knew nothing, to control giant industries, of which she had no knowledge.

The man in Bedroom F, Car No.13, was a lawyer who had said, 'Me? I'll find a way to get along under any political system.'

The man in Bedroom A, Car No.14, was a professor of philosophy who taught that there is no mind - *how do you know that the tunnel is dangerous?* - no reality - *how can you prove that the tunnel exists?* - no logic - *why do you claim that trains cannot move without motive power?* - no principles - *why should you be bound by the laws of cause and effect?* - no rights - *why shouldn't you attach men to their jobs by force?* - no morality - *what's moral about running a railroad?* - no absolutes - *what difference does it make to you whether you live or die anyway?*. He taught that we know nothing - *why oppose the orders of your superiors?* - that we can never be certain of anything - *how do you know you're right?* - that we must act on the expediency of the moment - *you don't want to risk your job do you?*

The man in Drawing Room B, Car No.15, was an heir who had inherited his fortune, and who had kept repeating, 'Why should Rearden be the only one permitted to manufacture Rearden Metal?'

The man in Bedroom A, Car no. 16, was a humanitarian who had said, 'The men of ability? I do not care what or if they are made to suffer. They must be penalized in order to support the incompetent. Frankly, I do not care whether this is just or not. I take pride in not caring to grant any justice to the able, where mercy to the needy is concerned.'

These passengers were awake; there was not a man aboard the train who did not share one or more of their ideas. As the train went into the tunnel, the flame of Wyatt's Torch was the last thing they saw on earth."

I think if we asked Rand about this she would roll her eyes, and say something about how they hadn't really been alive in the first place because they weren't thinking for themselves. I guess her argument is that for these people, life as they're living it is no better than death.

This doesn't do it for me morally, and even if you're a monster, it doesn't work out economically either. It's not only horrific, it's stupid. This enormous loss of life is also a huge loss of labor. The massive death tolls surely incurred by the breakdown of society would make it impossible to rebuild, especially since all their copper is now at the bottom of the ocean. How many people will be left in New York by the time the lights come back on? Who will the new trains they're planning serve when millions have already died because they were unable to get food or fuel?

The basics of comparative advantage also illustrate the problems with this strategy. Even if the unthinking masses were doing almost nothing before, the loss of their productive capacity would still be a hit to the overall economy, and the survivors will need to replace this lost work. While Dagny enjoys working as John Galt's maid/cook for a brief period, ultimately they will want other people to fill these roles if they're going to achieve their industrial dreams.

Birthing the system anew doesn't even ensure it will work out better in the long run. Since family is poor insurance of competence, the same process will likely repeat itself. How will this new world guard against just deteriorating back to a world of pull-peddling looters?

It gets even more depressing when we look at three characters who begin on the side of the looters, yet show promise to be more: Eddie Willers, Cherryl Brooks, and Rearden's Wet Nurse.

Eddie Willers is the very first character we meet. His parents and grandparents worked for the Taggarts before him, and he grew up spending summers with James and Dagny. He works in Dagny's office, and while he's no genius, he is generally competent. The last time we see him he has collapsed and is sobbing in front of a broken down engine, feeling "like the captain of an ocean liner in distress, who preferred to go down with his ship rather than be saved by a canoe of savages taunting him with the superiority of their craft." The other passengers have joined a roving caravan and while his status is ambiguous at the end, it's hard to imagine this ending any way other than his eventual death.

We know that there are some regular people (read: non-geniuses) who are invited to Galt's Gulch and allowed to live there. When Hank Rearden joins Galt's strike we know that Rearden's

secretary and several of his other employees get to go with him. But for whatever reason, Eddie doesn't make the cut.

Cherryl Brooks is the shoppirl who marries James. She only appears three times: her initial meeting with James at the dime store where she works, her wedding, and the final night of her life, which gets an entire chapter to itself (very unusual). Realizing all the things she had admired in James were actually achieved by Dagny, she goes to Dagny to apologize for misjudging her. Dagny attempts to make her feel better, but after they part, Cherryl wanders off through the streets of New York, slowly unravelling until she ultimately flings herself off a bridge to her death.

Rearden's "Wet Nurse" has a similar fate. His name is actually Tony, but this isn't revealed until the scene where--guess what!--he dies. He was originally sent by the government to keep a close eye on Rearden (hence the Wet Nurse nickname), but slowly comes to understand the evils of the people he works for and eventually begins to side with Rearden. He is injured helping protect the steel plants from the government agents trying to incite a strike and redeems himself before he dies in Rearden's arms.

Normally it's misleading to read too much into the trajectories of fictional characters, but *Atlas Shrugged* is a piece of propaganda in a way most novels are not, and its carefully choreographed plot is designed to give us lessons. All of these characters, who straddle the chasm between striker and looter ultimately realize the errors in their ways, and are redeemed as they...die?

There is no possibility of redemption. Nothing can be repaired or righted or restored, only pruned away.

VII.

When asked for the shortest possible summary, I told a friend that it's like Ayn Rand took too many amphetamines and asked "what if the capitalists went on strike" and then wrote 1,000+ pages about it. The things that people like about *Atlas Shrugged* (fun story with ~sexy times~, encourages self esteem, and validates your frustration) and don't like about it (long and rambly, seems like its advocating genocide by neglect) are all true.

I ended up in a weird situation where I found the first $\frac{3}{4}$ of every paragraph interesting and engaging only to wind up in the final $\frac{1}{4}$ being told that the answer is to be a huge jerk all the

time to everyone. Which...doesn't quite land for me. I also think this instinct to be cruel made her life far sadder and lonelier than it needed to be. Ultimately, Rand's interpersonal failings dampen her intellectual contributions.

The fantasy is based not only on direct and indirect violence--it also relies on a deeply American phantasm of inexhaustible natural resources that allows her to disregard the value that existing institutions and infrastructure hold. Maybe this read differently in 1957, but in 2020 it's more depressing than inspiring.

Today, I hope we can read *Atlas Shrugged* and take her advice to focus on reason, purpose, and self-esteem without feeling beholden to the more twisted aspects of her ideology. For now, I will work on finding a scenario in which I can appropriately compare someone's voice to mayonnaise.

Bad Religion: How We Became a Nation of Heretics

by Ross Douthat

Ross Douthat on Heresy, Christianity's Decline, and Possible Renewal

Are there present-day heresies, which might afflict you? (Yes, you!)

Why has Christianity declined in the United States? (Lots of reasons...)

What has replaced orthodox Christianity? ([the rise of the 'nones'](#), Pew Research Center)

Is there a defense for Christian orthodoxy today? (Yes.)

What should be the response of institutional Christianity to this decline? (some ideas at the end)

These questions (and more!) are asked and answered in Ross Douthat's ambitious book "*Bad Religion: How We Became a Nation of Heretics*" in 2012.

Important context: Douthat is a devout Roman Catholic—*Bad Religion* is not a diatribe against religion in the aggregate; instead, it documents the decline of orthodox Christianity since the 1960's (especially the decline of mainline Protestantism and non-immigrant Catholicism) and the rise of certain 'bad' religious substitutes (Douthat's adjective). To Douthat, these pseudo-Christianities are typified by Glenn Beck (Christian nationalism), Joel Osteen (prosperity gospel), Elizabeth Gilbert (the God Within/ 'spiritual but not religious'), and Dan Brown (choose your own Jesus) (descriptions Douthat's).

As evidence of both the decline of Christianity's importance to American society and the internal debate happening within Christianity, the United Methodist Church (my church, the second largest Protestant sect and one of the major mainline Protestant churches) is currently in a [schism](#) and nobody even knows.

In Douthat's wildest dreams, *Bad Religion* serves as the inspiration for detached Christians to re-gain their faith, and for secular people to perform self-reflection about their own dogmas. If there were one book I could make every Christian read, it would be this one.

But don't flee, non-Christians! The heresies that Douthat describe also implicate secular society. His chapter on American nationalism is prophetic and practically perfect, while his criticism of the prosperity gospel should have wide-ranging implications. And as for Elizabeth Gilbert, Dan Brown, and even Oprah? Douthat's not sparing any punches. *Bad Religion*, if read correctly, should challenge everyone, no matter your religion or political beliefs. It is not a perfect book, though; the "Accommodation" chapter about theological liberalism is inadequate.

But in this, I concur with Mathew Yglesias: “And critically, by ‘highly recommend it to you’ I do not mean ‘I agree with all the takes.’...It’s important to read strong writers with big, influential ideas and understand what they’re saying.” ([link](#))

Douthat’s definition of a heretic is someone who rejects the synthesis and complexity inherent in Christian orthodoxy. To Douthat, Heretics make a false choice to simplify Christianity. These ‘solutions,’ while seemingly justifiable or logical, often result in a corrupted faith. Heresy sets out to be simple, rational, and appealing, but it often ends up being extreme and misguided.

The Heresy of Christian Nationalism

In only 30 pages, Douthat explains the intended meaning of “city on a hill;” the problem with American founder worship; the relationship between Glenn Beck’s political rants and theological Americanism; one appeal of blood and soil nationalism; the potential problems of American exceptionalism; the healthy patriotism of George Washington, Abraham Lincoln, and Dwight Eisenhower due to humility; the messianism of Woodrow Wilson, John F. Kennedy, Lyndon B. Johnson, and George W. Bush; the ouroboros of messianism and apocalypticism; and the consequences of Christianity’s inability to transcend partisan politics.

One can hear the ring of “Make America Great Again” in American Nationalism’s “call(s) to restore the country’s greatness” (244, *Bad Religion*). If nothing else, you should buy this book, read this chapter, and evangelize to the masses. Just casually say “Douthat was right about the heresy of American Christian nationalism, and its pernicious effects in American society. The twin terrors of messianism and apocalypticism have and will continue to wreak havoc in America and in the world.”

Here’s just a sample of the glory that is this chapter (263, Douthat, *Bad Religion*):

But if messianism has done more good than apocalypticism, it has probably done more damage as well. Precisely because the messianic style has been more influential among the American elite, the consequences of messianic excess have generally been more comprehensively disastrous. Apocalypticism is rarely harmless, but its very marginalization limits its destructive power. Witch hunts are dangers and deadly, to be sure. But “wars for righteousness” often have far more victims, and they do more lasting harm.

This is not how our history is usually remembered. But far more people suffered, at home and abroad, because messianic chief executives like Wilson and Teddy Roosevelt took the country repeatedly to war than suffered because of the excesses of Protestant fundamentalism in the same era. (*Inherit the Wind* bestrides high school reading lists, but the Roosevelt-championed Filipino War,

America's first exercise in disastrous nation-building, has been more or less erased from the national memory.

While there's much, much more, let's move on. As I said, the book is ambitious.

The Heresy of "Pray and Grow Rich"

The infamous "Prosperity Gospel" has its origins in the New Thought movement, Pentecostalism, and an American strain of pro-capitalism. According to [Did Christianity Cause the Crash?](#) by Hanna Rosin, approximately 50 of the 260 largest churches preach the prosperity gospel today.

The prosperity gospel makes Christianity purely transactional: if you believe in Christ, you will be rewarded handsomely in this life.

This crass materialism is revealed when Osteen reflexively cries "No big deal, God's promised me double," when Joyce Meyer says "obey me and do what I tell you to do, and you'll be blessed," when Kenneth Hagin exhorts "name it and claim it," when Kenneth Copeland promises "Do you want a hundredfold return on your money? Give and let God multiply it back to you. No bank in the world offers this kind of return! Praise the Lord!" (Osteen at 189 in *Bad Religion*, Meyer at 188 in *Bad Religion*, Hagin at 188 in *Bad Religion*, Copeland at Tara Burton's [The Prosperity gospel, explained: Why Joel Osteen believes that prayer can make you rich](#))

In light of this theology, it is unsurprising how many of these preachers own private jets or even commit [fraud](#).

Beyond the obvious criticisms that the prosperity gospel is incomplete (Douthat's quip "Every day is Friday, but there's never a Good Friday" is illustrative), the core problem is that the prosperity gospel implies that if you're not successful, it's because you're not a "real" Christian. From pastor Fernando Garay: "Ten Christians will say that God told them to buy a house. In nine of the cases, it will go bad. The 10th one is the real Christian." (209, *Bad religion*)

This fallacy, though, is not limited to the prosperity gospel:

1. Fictional Steve Prefontaine in movie "Without Limits" when he claims "I run to see who has the most guts."
2. The successful science professor who says "Whenever I worked hard, the experiment worked out for me, so you must be slacking." (not a paraphrase...)
3. "The people who think admission to an Ivy League university was because they were superior, intellectually and perhaps morally, to those not as fortunate. They wanted to be the college admission version of John Calvin's 'elect,' and they didn't want luck to play any part of their admission" from [Ethical College Admissions: Ivy Lottery](#)

- a. Note: this is true even if elite colleges stopped admitting legacies or athletes or stopped preferentially accepting people willing to pay full tuition (i.e. early decision)
 - b. The truth is almost everyone who is accepted to an elite college is qualified and deserves to be there (yes, the athletes and legacies too) — it's just that there are many, many other people who are also qualified.
 - c. It's why I support a lottery like Nathan Robinson's [Admit Everybody](#)
4. the entrepreneur who thinks his business succeeded only because he worked hard

Douthat seems to be worried that “the result is a country where religion actively promotes the sort of recklessness that produced our economic meltdown, rather than serving as a brake on materialism and a rebuke to avarice” (5, *Bad Religion*).

I'm more worried that the prosperity gospel (and related fallacies) lead to a justification of the status quo. I still want everyone to work hard and strive for perfection, but if you lose the “chosen elect” assumption, we might have the following results: scientists could write papers, and earn respect, for experiments that didn't work. The sports world could stop reflexively calling every first round pick a “bust” if they fail to become all-stars. In general, successful people would gain humility. From David Kinney's [The mathematical case against blaming people for their misfortune](#), “this leads to the conclusion that compassion, not blame, is the appropriate attitude towards those who act in good faith but whose bets in life don't pay off.”

The Heresy of the God Within

This is perhaps Douthat's bravest chapter, where he takes on the “spiritual but not religious” creed, or what he terms, “the God within.” Among those implicated include Elizabeth Gilbert, Deepak Chopra, Marianne Williamson, Oprah Winfrey, Ralph Waldo Emerson, and the metaphysics of both Disney cartoons and George Lucas's Jedi.

Douthat observes that Gilbert, as recounted in her confessional, *Eat, Pray, Love*, does not change her theology at all even after she had all of these “earth-shaking, all-enveloping encounters” with God (214, *Bad Religion*).

As for Oprah? Well, the subtext of New Thought that runs through the prosperity gospel also suffused her show. To Douthat, “You get a car” if not a reward exactly, seems to always follow the proclamation of “following your spirit.”

An keen reader might recognize the intriguing placement of the words “I” and “fascinations” in Douthat's phrasing about Gilbert's journey: “an ample publisher's advance allowed her (Gilbert) to unite her three newfound “I” fascinations—Italy, India, Indonesia—into a world tour and then a travel book” (213, *Bad Religion*).

And thus we learn Douthat's criticism of the God Within: that it leads inexorably to pride, vanity, adultery, greed, and gluttony.

"Do-it-yourself religion" falls prey to the problem that if you're not careful, it is very easy to justify your vices. "Hakuna Matata" and "Bare Necessities" are great songs, but are rather insufficient anthems. Douthat recommends the necessity to "imitate Christ," but regardless, there has to be an element of responsibility, sacrifice, and cultivation of ethics in order to achieve eudaimonia.

The Heresy of "Brave Certainty"

Here Douthat examines some academic scholars' "brave certainty" in the search for the "real" Jesus. Douthat thinks this quest is really an excuse to eliminate the parts of Jesus that are paradoxical or confusing. If taken to the extreme, some practitioners essentially form a Jesus of their choosing. Douthat also thinks this "quest for the historical Jesus" is emblematic of liberal Christianity's decline and, perhaps, the original sin of our other heresies. (172, *Bad Religion*).

And Douthat thinks it's all Dan Brown's fault. So let's talk.

I agree with Douthat that you shouldn't read the *Da Vinci Code* to get the true history of Christianity. The Gnostics did not portray Jesus as purely human, the decision of which books to include in the New Testament is very justifiable, and the Priory of Sion probably isn't real. ([The Da Vinci Code](#), Wikipedia)

But Douthat also distrusts Brown's "theological project.". Granted, I might feel differently if I were Catholic, but I enjoy Brown's books and basically agree with Brown's universalism.

But before we answer the question of "Why Christianity Declined?," let's take a small detour and reflect on this "brave certainty" that so irritated Douthat.

A statement can be objectively wrong if it does not admit the possibility of error, or if it is oblivious to the fact that a statement can be true, false, unknown, or unknowable.

In one study, Philip Tetlock shows that many people commonly discover that half or more of their 90% confidence ranges fail to contain the true answer (Tetlock, [Superforecasting: How to upgrade Your Company's Judgment](#), Harvard Business Review). In other words, most people's confidence calibration is off. Here are some examples of this phenomenon:

Christian Statements

1. “The stories (New Testament)...tell us nothing whatsoever about the origins of Christian faith but quite a lot about the origins of Christian authority” John Crossan (171, Bad Religion)
2. The “exodus story” is a myth because we do not have archaeological evidence.

Science

1. No aliens exist.
2. One study shows flossing doesn’t reduce cavities or gum disease, so flossing has no benefit.
3. Hunter-gatherer societies never had woman hunters.
4. The Out of Africa hypothesis is true because all ancient hominid were fossils found in Africa.
5. “While mammals and birds possess the prerequisite neural architecture for phenomenal consciousness, it is concluded that fish lack these essential characteristics and hence do not feel pain.” a 2014 [scientific article](#) published in Springer: Biology and Philosophy

Polling (aka Silver was right and I will not be taking questions about the 2016 election at this time)

1. “I get why Silver wants to hedge. It’s not easy to sit here and tell you that Clinton has a 98% chance of winning. Everything inside us screams out that life is too full of uncertainty, that being so sure is fantasy. But that’s what the numbers say.” (Ryan Grim, [Nate Silver is Unskewing Polls –All of Them—In Trump’s Direction](#))
2. While 538 is great and mostly well-calibrated, they need to update their NBA playoff predictions. Their RAPTOR model’s skepticism of the Lakers wasn’t credible (the model actually favored the Rockets and Heat in last year’s playoffs)

Covid-19 Edition

1. I haven’t had Covid-19 symptoms, so I haven’t had Covid-19.
2. “Is this going to be a deadly pandemic? No.” ([deleted Vox tweet](#)).
3. “[The Conspiracy theories about the origins of the coronavirus, debunked](#)”
 - a. “But already, virologists who’ve parsed the genome and infectious disease experts who study coronaviruses say they have enough evidence the virus is brand new and came from nature” ([fifth paragraph](#), vox, again)
4. “[baseless theory that Covid-19 was manufactured in a Chinese Lab](#)” from NYT

Note Vox’s article, “[Covid-19 vaccine trials are showing promising results. A lot can still go wrong](#)” is epistemically correct.

I’m susceptible too

1. My “family myth” about my ancestor who changed his name from Andersson while in transit from Germany is false. (I “knew” this was so because there was no evidence, other family stories are also false, and the commonality of “Andersson” was suspicious). Turns out, I’m related to an Anders Larsson from Sweden. Given Swedish patronymics, this means my ancestor originally called himself Andersson.

But don't overcorrect. Don't be a [scared, half-committed, appeasement-seeking poll herder](#). (said in [this tone](#))

The problem with poll-herding or being a scaredy-cat is that poll herding can “make the *average poll* more accurate even as it makes the *polling average* worse. (For economics nerds — this is sort of a [tragedy of the commons](#) problem.)” (From [538](#))

Instead, be like [Ann Selzer](#). Do the work, then let the chips fall where they may.

Why is Christianity in decline?

To setup Douthat's desired framing of a “fall,” he first presents condensed biographies of Ronald Niebuhr (Protestant), Billy Graham (Evangelical), Fulton Sheen (Catholic) and Rev. Martin Luther King Jr. (African American Baptist) as touchstones to provide a story of Christian revival in the 1950's and 1960's. In his telling, King's “Letter From a Birmingham Jail” drew some of its moral force from centuries of Christian tradition (King even quotes Niebuhr).

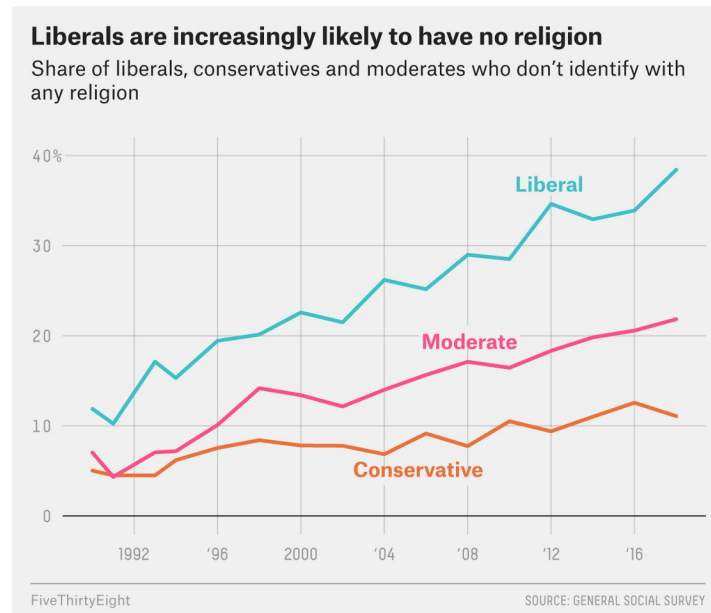
Notably, this Christian revival really happened—church attendance per week was approximately 60% of adults in the 1960's while it's less than 30% today. Though if you broaden your perspective, picking 1960 as a starting point is somewhat arbitrary. Lyman Stone's report: “[Promise and peril: The history of American Religiosity and its recent decline](#)” concludes:

“According to membership data, religiosity in America peaked sometime between 1940 and 1970. Religious membership *rose dramatically* during and after World War II in particular. But for the past 50 years, religious membership has been *in decline*.” (13, *Promise and Peril*, italics mine)

Nevertheless, the decline is real. America is a big country, and religion is complex and personal, so there just isn't one reason for the decline. Douthat suggests five causes: political polarization, the sexual revolution, an increasing global perspective, an increase of wealth, and the elite class's dismissal of Christianity. Christian Smith, a religion professor at Notre Dame, suggests that the decline is caused primarily by the “the association of the Republican Party with the Christian right, the end of the Cold War, and 9/11” (Derek Thompson, “[Three Decades Ago, America Lost Its Religion, Why?](#)”). Others have suggested the Catholic Church child sexual abuse scandal, the general fracturing of society described in Robert Putnam's [Bowling Alone](#), family instability, or even “adulthood.” (Thompson) Stone would add a possible snowball effect with respect to decreasing “club goods” and the decline in marriage rates (42,45, *Promise and Peril*).

Some theories that are probably false include the rise of education, urban life, industrialization, or worship style (i.e. rock music vs. traditional hymns). (Stone, *Promise and Peril*).

But if you look at the data, it's clear that liberals are becoming non-religious at a much higher rate than moderates or conservatives. This article at 538, [The Christian Right Is Helping Drive Liberals Away From Religion](#)” by Amelia Thomson-Deveaux is very persuasive and provides evidence for the claim that the biggest reason for the decline of Christianity is the Christian right.



We've talked a lot about political polarization with respect to political parties and ideological conformity, and I'm not convinced that political polarization is a bad thing for political parties or the American political system. But the association of Christianity with the Right is very bad for Christianity writ large. Ideally, a Christian church should transcend partisanship and sometimes challenge the political consensus of the Republican Party, the Democratic Party, or both. Jesse Jackson, Al Sharpton, Pat Robertson, and Mike Huckabee running for president are symptoms of too many religious leaders (Robert Jeffress, Jerry Falwell Jr. among others) becoming overtly partisan. More obviously, the evangelical community identifying with the Republican party "magnified their political strength even as it compromised their moral credibility." (69, *Bad Religion*). To Douhat, it's depressing that "frequent churchgoers were the most fervent supporters of waterboarding detainees, among other seemingly un-Christian practices" (141, *Bad Religion*). As Douhat observes about the Christian Right: "resentment, partisanship, hero worship, martial language...these were unattractive features for a movement whose primary goal was supposed to be winning converts, not elections" (141, *Bad Religion*).

So how does this phenomenon work in practice? Let me give you an example.

After earning my Eagle Scout in 2009, the Boy Scout Council gave me a signed copy of a book by Gov. Rick Perry called: *On My Honor: Why the American Values of the Boy Scouts Are Worth Fighting For*, with blurbs by Sean Hannity and Newt Gingrich. The

book attempted (rather unpersuasively) to defend the Boy Scouts' ban on homosexual or atheist scouts/scoutmasters. "Traditional" American and Christian values were used to defend this ban against the "radical leftists" and the "secular humanists." Oh, the horror. Later, Perry dismissively writes: "For that matter, they could, if they wished to form a national atheist youth group" (183, Perry, *On My Honor*). Given this "my way or the highway mentality" of many defenders of "traditional" Christianity (the 2019 United Methodist Special Called General Council is one example of many), and the increasing stakes of conflating Christianity with the Republican Party, it is perhaps not surprising that many liberals chose the door.

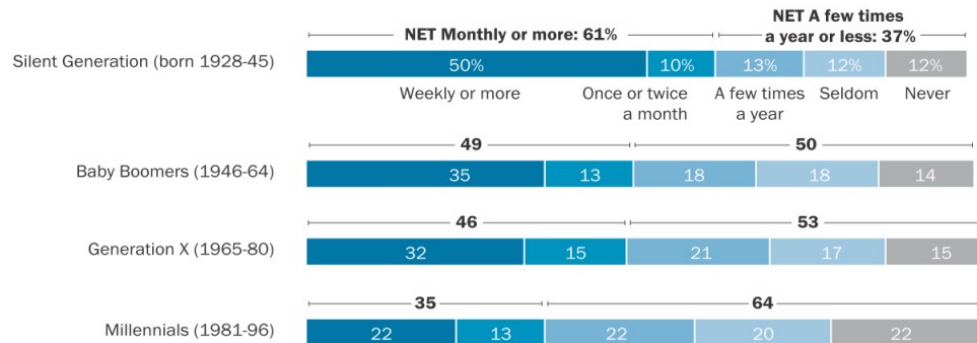
To evaluate the other causes, though, it's important to have the correct mental model about how large and widespread this decline actually is. This graph from Pew Research Center shows that the decline in religiosity is largely a generational effect.

Large generation gap in American religion

In 2018/2019, % of U.S. adults who identify as ...



In 2018/2019, % of U.S. adults who say they attend religious services ...



Note: Don't know/refused not shown.
 Source: Aggregated Pew Research Center political surveys conducted January 2018-July 2019 on the telephone.
 "In U.S., Decline of Christianity Continues at Rapid Pace"

PEW RESEARCH CENTER

And before you think this phenomenon is just because young people are more likely to be less-religious; it's pretty clear that the generational difference in religiosity is not just an age effect. In fact, every generation has declined in religiosity over the last 20 years. (table from Gallup's [U.S. Church Membership Down Sharply in Past Two Decades](#)) So people are not even attending church more as they age.

Changes in Church Membership by Generation, Over Time

	1998-2000	2016-2018	Change
	%	%	pct. pts.
Traditionalists (born in 1945 or before)	77	68	-9
Baby boomers (born 1946-1964)	67	57	-10
Generation X (born 1965-1979)	62	54	-8
Millennials (born 1980-2000)	n/a	42	n/a

Note: Given that Gallup's polls are based on the 18+ U.S. adult population, the 1998-2000 time period would have included only a small proportion of the millennial generation.

GALLUP

Not only are people not getting more religious, but younger generations are less religious than older generations. From the table, we can see that at the same age, 62% of Gen X belonged to a church while only 42% of Millennials do. The key insight is that the religious preferences of the next generation start young:

“Rather, they (Raphaël Franck and Iannaccone) found that religiosity is determined early in life: Kids raised without religion tend to become nonreligious adults, and vice versa. The vast majority of change in religiosity over time in most countries is not because adults converted away from a given religion, but simply because the next generation was less religious ... In other words, the story of secularization in America is *not* mostly a story of lots of people who were raised religious leaving their religious faith as adults. It is a story of fewer people having a religious upbringing at all. (42, *Promise and Peril*)

Being religious is a habit, so if we want to look at the most important drivers of why America is

becoming less religious, then the answer will be found in why kids aren't growing up religious.

So this leads us to the second biggest reason for Christianity's collapse: secular education. This point is related to Douthat's use of the word “class. Specifically, the fact that almost every elite institution would only talk about Christianity in order to reject it—and then, to simply dismiss it. If all you learn about Christianity from school is the Crusades, the Spanish Inquisition, the Catholic Church's opposition to Copernican Theory, *The Scarlet Letter*, the Scopes Monkey Trial, and Voltaire's *Candide*—Christianity doesn't look very appealing. If your family only attends church a few times a year, many of these kids will grow to become “nones.”

Douthat also blames the birth control pill and contraceptives more broadly. As more people reject Christian sexual ethics, it becomes easier to simply leave the flock. According to Douthat, “separating sex from procreation more completely than any previous technology, the birth control pill also severed the cultural connection between Christian ethics and American common sense. For the first time in human history, it was possible for the poor and middle class as well as the rich to imagine being *safely* promiscuous.” (71, *Bad Religion*). Today, though, it seems that Christianity's response to LGBTQ people will be more influential.

Douthat also blames an increasing global perspective and conflation of Christianity with colonialism and racism. In essence, he thinks a cosmopolitanism and white guilt, along with the rise of relativism, presented a challenge to Christianity. I'm skeptical. If you live in a society with multiple religions and denominations, then any one religion has the problem of not having an "exclusive truth claim."

Next, Douthat blames increasing wealth. The increasing wealth of American society meant that the opportunity cost of becoming a minister increased: the average salary of a protestant minister with a graduate degree only increased 10% between 1970-1990, while salaries of, say, a doctor or lawyer rose 37%. (79, Bad Religion) (Note, this fact is clearly not true of some megachurch pastors, but Douthat and I think this phenomenon is not a satisfactory solution). This increase in wealth was bad for two reasons: many people who might've become ministers, fathers, nuns, or preachers instead became something else. But also, increasing wealth and increase in movement made it harder to create a community. This may contribute to the difficulty of recruiting talented ministers, which thus may make it more difficult to have a strong Church community.

The Catholic sexual abuse crisis is a huge cause. An astounding [10% of Americans](#) are former or non-practicing Catholics. And while it's hard to know the counter-factual, it's clear that the Scandal caused a crisis of faith. (And not just for Catholics. There is sexual abuse crisis in Protest denominations as well, and the Catholic abuse scandal affects Christianity in the aggregate because the Roman Catholic Church is the largest Christian organization).

The catch-all of "family formation" also probably has an effect, but it's pretty hard to know just how much this has or hasn't had on religiosity. Stone adds that once the decline starts, there is a possible snowball effect. Organized religion provides fewer "club goods," while it becomes easier and easier for people to meet and live with other non-religious people.

Who are the "Nones?"

Well they're not all atheists—only 4% of American adults self-identify as Atheists. Agnostics make up another 5%, while 17% describe their religion as "nothing in particular" ([Pew Research Center](#)). So together, the religiously unaffiliated is 1/4th of America.

Somewhat surprisingly, only 10% of Americans say they don't believe in God ([Pew Research Center](#))

But while only 25% of the population is religiously unaffiliated, a substantial 40% of Millennials are. If I were a betting man, I would bet that the decline of Christianity will continue apace.

What is Douthat's Defense of Orthodox Christianity?

Orthodoxy obviously includes "Christ's incarnation and atonement, the Trinity and the Virgin Birth, the forgiveness of sins and the possibility of everlasting life, ... a rejection of violence and cruelty, a deep suspicion of worldly wealth and power, and a heavy stress on chastity "(10, *Bad Religion*).

Beyond the basic tenets of the Christian faith, Douthat beautifully proposes that at its best, orthodox Christianity contains a commitment to mystery and paradox. These riddles include Christ being human and divine simultaneously, God being omnipotent yet allowing free will, living in a World corrupted by original sin yet fundamentally good, how the God of the Old Testament is the same as the God of the New Testament, how God forgives every sin yet still sets impossible moral standards, an insistence that faith alone is will save you, but that faith without works is dead, and that Christianity properly conceived has a role for everyone-the family, the celibate, the politician, beggars, soldiers, pacifists, criminals ...

As for the Christian teaching on sex and the family? Well, [family structure matters](#). And the reappraisal of sexual ethics in the #MeToo era should be an opportunity for Christianity to regain some cultural cachet.

And while I understand the financial benefit and social convenience to premarital cohabitation, cohabitation is still [correlated](#) with divorce. Even an article that largely defended the practice of cohabitation [admits](#) that cohabitation can often prevent people who should break-up from actually breaking up. And cohabitation is also not good for children. According to Bradford Wilcox, a Professor of Sociology at the University of Virginia:

"Cohabitation fosters enough intimacy to facilitate childbearing but not enough commitment to make people deliberate about their choices to become parents. The result, an unplanned birth, can pose real problems to their relationship and to their future odds of successfully marrying" (Lauren Fox, [The Science of Cohabitation: A Step Toward Marriage, Not a Rebellion](#)).

But let's talk about "Accommodation."

Bad Religion's third chapter, "Accommodation" is not great. In essence, it's a repetition of Dean Kelley's 1973 "Why Conservative Churches Are Growing" which is used almost exclusively by the traditional Christianity faction as a club against the theological liberals. To the traditional Christians, "accommodation" is tantamount to an indictment. Basically, Douthat believes that the Church has two choices: resistance or accommodation. He of course associates "accommodation" with Teilhard Chardin, Harvey Cox, and Bishop James Pike; and he partially blames "accommodation" as responsible for Christianity's decline.

His fourth chapter, “Resistance,” is an apology for “traditional Christianity” and the rapprochement between Evangelicals and Catholics. And it’s mostly fine, I don’t dislike John Paul II (though in hindsight, John Paul II’s response to the Sexual Abuse Crisis seems woefully deficient). And I have respect for the evangelicals Jimmy Carter and George W. Bush. Though Douthat does annoyingly imply that the Lutheran John Neuhaus (who founded the influential religious journal *First Things*) among other prominent protestants made the correct decision to convert to the Roman Catholic Church. After reading Douthat’s *To Change the Church*, I kinda want Douthat to justify why he hasn’t converted to conservative Lutheranism. I thought that Catholics believed in Papal Infallibility.

But seriously, Douthat seems to misunderstand the unease faced by many theological liberals. When I read Andrew Sullivan’s earnest plea in [Here Comes The Groom](#), my initial response was that I *knew* in a Platonic sense that Sullivan’s cause is just. I still don’t think the “traditional Christianity” response to Sullivan’s argument is very good, but Douthat is right about one thing: I don’t know if the quality of their argument would’ve mattered to me anyways. I knew that Sullivan was right, and I didn’t need to read the Bible to know that. This tension, of course, is also found in the Euthyphro Dilemma.

But my frustration with the “traditional Christianity” faction goes deeper. One proponent of “traditional” Christianity wrote a twitter thread about the Methodist schism that was especially illustrative and infuriating. Among other things, he accuses white theological liberal American Methodists of being racist:



The debate, as everyone knows, was about whether the United Methodist Church would sanction gay marriage. The fact that this flashpoint also fractured the church politically made the schism possible. So while I’m sad that there’s a schism, it might be for the best.

I don’t expect everyone to want to be Methodist. But Methodism has open communion, and it is important to me that anyone *could* be Methodist. And at least the UMC will allow conservative Churches to keep their property—for the most part, it’s an amicable split.

Can Christianity Recover?

Douthat suggests four possibilities for Christianity’s recovery.

1. **The Postmodern Opportunity.** Here, Douthat suggests that it might be possible for people to re-gain the spirit of Christianity in the first-century. The scholarly “radical Orthodoxy” or evangelical “emergent” movement would re-build Christianity from the ground-up with house churches and small groups; with prayer and storytelling.
2. **The Benedict Option.** Popularized by [Rod Dreher](#) and practiced by the Church of Latter Day Saints, here Christian communities would self-segregate in a “mustard-seed” strategy.
3. **New Christendom.** From Philip Jenkins’s [book](#). In this scenario, the emerging centers of Christendom of Sub-Sahara Africa, Pacific Asia, and Latin-America re-evangelize North America and Europe.
4. **Age of Diminished Expectations.** Last, Douthat suggests that cultural crises might cause a reassessment of modern life and drive people back to the faith of their forefathers.

Personally, I am partial to the postmodern opportunity (though I would be hesitant about making the analogy too strong to the early Christian church). I respect the confidence of the Benedict option, but in practice I am sad that the Mormons left the Boy Scouts. As for New Christendom? American Christianity is one of the best avenues we have to assimilate the many immigrants needed to achieve [One Billion Americans](#).

As for Douthat’s fourth suggestion, well, [I’m a dreamer too](#).

But how should Christianity respond?

Douthat provides four suggestions.

1. **political without being partisan.** It may be the case that we have to have liberal and conservative church denominations, with the exception of the Catholic Church. But just because you attend a liberal or conservative church, it doesn’t have to be a partisan church.
2. **ecumenical but also confessional** Douthat suggests that each denomination should conserve its inheritance. Catholics should still be robust Catholics. Methodists should continue to be inspired by John Wesley. Every denomination should remember its own traditions and inheritance.
3. **moralistic but also holistic** Douthat suggests that Christianity must not abandon criticism of the seven deadly sins, or to forget the importance of chastity and fidelity. That Christianity must always remember St. Basil’s exhortation that “The bread that you possess belongs to the hungry.”
4. **oriented toward sanctity and beauty** Douthat suggests Christianity requires “public examples of holiness, and public demonstrations of what the imitation of Christ can mean for a fallen world” (292, *Bad Religion*). If this revival is

accompanied by great Christian songs, art, literary works, or architecture so much the better.

What does this look like practice?

It looks a lot like David French on [decency](#). Or Elizabeth Bruenig on [forgiveness and atonement](#).

At the end of his report, Stone recommends that Christians support liberalized zoning rules to make it easier to make new churches, make it easier for the creation of new parochial schools, support child allowances to support childrearing, and a removal of the marriage penalty in tax code and welfare law. (54, *Promise and Peril*).

These are all good suggestions, but I have more.

Christians should not support the Boy Scouts ban on atheism, and instead should support the Boy Scouts adding a Utilitarian, Stoic, Epicurean, and Effective Altruism religious emblem. Some Christians claim that “You must believe in God to be moral” is not true, nor is its corollary. Instead, Christians should claim that “Believing in God in a church helps me be moral.” Christians could model behavior and, ultimately, find substantive agreement about shared goals.

Since secular education is important, I would stop trying to ban Harry Potter (or any book) and instead try and install Quaker abolitionist writings and Catholic opposition to eugenics into the school curriculum. If “pro-life” in the context of abortion also means “pro-life” in the context of eugenics, the Catholic teaching makes more sense. The contribution of Christian churches by building hospitals, schools, orphanages, and universities shouldn’t be forgotten.

Long term, there should be an unofficial “Third Testament” that, at minimum, includes Rev. Martin Luther King Jr.’s “Letter From a Birmingham Jail.” MLK’s Letter has obvious analogues to a Pauline epistle, and this heritage can and should be made explicit. This new canon could also include St. Augustine’s *Confessions*, Dante’s *Divine Comedy*, G.K. Chesterton’s *Father Brown*, or any number of writings from C.S. Lewis.

And What About the Nones?

Douthat’s go-to gotcha of “deed, not creeds” he uses against the theological liberals is annoying. For one thing, it’s not true about the Protestant Mainline. And second, to the degree that it is true, I want those people to also have a productive community.

As for those liberals who felt pushed away from the church? Please, don’t just accept Democratic Party politics as an adequate religious substitute.

From Derek Thompson’s “Three Decades Ago, America Lost Its Religion. Why?”:

Making friends as an adult without a weekly congregation is hard. Establishing a weekend routine to soothe Sunday-afternoon nerves is hard. Reconciling the overwhelming sense of life's importance with the universe's ostensible indifference to human suffering is hard.

And it is hard. I rarely lace up my running shoes since my Cross Country days, and I haven't even touched my trumpet since the last day of Marching Band. Going alone is hard, and institutions or communities of people with shared goals make everything easier. But no-one said re-knitting the fabric of American society would be easy.

If you believe in "deed, not creeds," you can join [Unitarian Universalism](#) or even [Effective Altruism](#). If you're just generally agnostic, you can start or attend a [church of Freethought](#).

Inspired by Benjamin Franklin, you could start a [Leather Apron Club](#) for mutual improvement. All you have to do is meet every Friday to debate morals, politics, and natural philosophy. It can even start in a tavern.

Maybe no one will show up, but maybe they will. Maybe they're just waiting for someone else to take that first step.

That someone else?
[It could be you.](#)

Behave by Robert Sapolsky

I learned two things from this book - one from the first half, and one from the second. The first, Robert Sapolsky wanted me to learn. The second - I don't think so.

But first - a quick introduction. *Behave* describes "the biology of humans at our best and worst". More specifically, it investigates the biology of aggression and altruism, from neurons to psychology to culture. Many of the books chapters are laid out like circles surrounding an action - for example, reaching out to touch someone's shoulder, or pulling a trigger. Sapolsky reminds us that the motivation for the same action can be good or evil - a hand reaching out to offer compassion, or a hand reaching out to shove someone. A finger on the trigger in a genocide, or pulling the trigger to defend an innocent bystander.

We start by looking at the seconds before the act - the realm of the brain. We dive into neurons and action potentials, and understand how a decision in the brain results in a physical movement. We zoom out to a few minutes, and see various parts of the brain, and how they were involved (learning along the way how scientists discovered these connections). We zoom out further - hours or days before, and think about the effects of hormones - oxytocin, adrenaline, testosterone. Zoom out further - the effects of childhood and parenting, and of our genes. Further - the effects of culture. Further - the evolution of human beings.

If this sounds interesting, then I recommend this book. Fair warning - although the book presumes little knowledge of biology (there are appendices with introductions to neuroscience, endocrinology, and proteins), once a concept is introduced, Sapolsky expects you to remember it. I have to admit, I have no idea what region of the brain the vmPFC is, or what significance it has if it's activated. But it seems to be important because it activates at least once every couple pages in this book.

And that leads us nicely into the first lesson:

Everything is connected

A particular neurotransmitter is released, and it results in lowering the activation threshold for some associated neurons. Except - in the presence of a different neurotransmitter, it has the opposite effect. Except - the receptors on the neuron can affect how the neurotransmitter is interpreted.

And then it gets even worse, because that same neurotransmitter also serves as a step in generating a hormone. So the short term effect might be to increase brain activation in a region, but in the long term, the hormone decreases activity in that region.

And everything is like this, and they're all related to each other, up and down the levels, around and around in circles. Childhood trauma affects the genes through epigenetics, and the neurons through stress hormones, and the connections between brain regions since different regions become more important... and then each of these affects our behavior, which then affects our brain, genes, body, hormones...

To recap - any particular external influence (external to the system at a particular level - it might be a cell, an organ, a body, a person) has effects that depend on context - like our neurotransmitter in the first example. And that external influence will have short- and long- term effects that change how the next influence will be interpreted.

But saying that "everything is connected" is equally informative as saying "nothing is connected". The true value lies in discovering which connections are important and causal, and which are not. And this book does a great job of laying out some important connections, and how they were discovered. But this leads to the second thing I learned...

Biology is too hard

Even in the chapters focusing only on interactions among brain regions, I recognized that Sapolsky was sometimes overstating his case (although in general, he is very careful to specify how confident we should be in any given statement). But then we get to the later chapters, where we take the biological context that we built up in the first half and apply it to the big questions - what makes us moral? why and how do we form hierarchies? what allows aggression and altruism on a societal scale? - that the science falls apart.

Most of the interesting results that Sapolsky cites - that certain cultures have tendencies towards cooperation or competition, that unclean environments make us more cautious towards outsiders, that a particular action can increase our altruism - he depends on priming studies. In the typical case, a word or image is quickly shown to the subject - for example, either the word "doctor". Then that person will recognize the word "nurse" faster than an unrelated word. An interesting effect - and one that has been almost entirely a [casualty of the replication crisis](#).

This doesn't mean the conclusions Sapolsky draws are wrong - it could be that priming studies are just easy to write about, or are the most accessible to the non-biologist audience. But it doesn't give me much confidence that anything in the last half of the book is evidence of... anything. When I read it, I tried to remind myself of that fact every page. I even wondered if I should stop reading, so as not to contaminate myself by thinking I'm receiving evidence when I'm really not. But in the end I persevered. If I were you - I wouldn't. The last half of the book is fascinating and gives you a new perspective on many important questions. And that's why it's dangerous - because that new perspective is founded on faulty studies, giving us no reason to privilege it over the boring old perspective we already have. You should only read it so that you can identify

the things you already believe that are mentioned in this book, and then decrease your confidence in those beliefs.

I joke - it's not that bad. But it makes me realize how difficult this all is. In his writing, Sapolsky comes across as thoughtful, meticulous, not afraid of controversy, but also not drawn to shocking statements only for the sake of contrarianism. He appreciates nuance, and is incredibly intelligent. And still - he was unable to find the truth.

So be careful when it comes to biology and especially social science - hesitate to throw out your own observations for the sake of science.

Brave New World by Aldous Huxley, Or, Should I wirehead my cat?

We start at the beginning, with babies growing in bottles to a civilization evolved past the need for families. This is Aldous Huxley's *Brave New World*, the seminal work of the utopian genre. Here, the motto of the World State is "Community, Identity, Stability". Huxley paints a bleak picture of technologically enabled hedonistic excess. Unfortunately for possibly all of humanity, *Brave New World* is much more popular than its utopian counterpart *Island*, written thirty years later. The attention grabbing portrayal of a society that values happiness to the exclusion of culture, freedom and even love has tainted the idea of paradise-engineering in the public eye for decades. Taking into account the immense amount of human and animal suffering in the world, stymying research into abolishing suffering is a moral catastrophe, one Huxley contributed to.

Brave new world conflates paradise engineering (using technology to fulfill hedonistic imperatives, ie make everyone happy all the time) with wireheading. Wireheading is changing your reward architecture to be happy all the time. The term refers to experiments where wires were used to electrically stimulate regions of a rat's brain, short circuiting the brain's reward mechanisms. This is endless reward, without the tolerance or satiation of sex or drugs. Rats will push a lever that provides inter-cranial stimulation until they die of dehydration. Humans find wireheading repugnant, at odds with their most sacred values. In *Brave New World*, humanity has wireheaded not a rat but all of humanity. With children freed from the tyranny of families, hypnotic conditioning to accept your place in society, and endless soma, the hangover free high, "everybody's happy nowadays". (BNW 96)

Brave New World functions as a piece of propaganda, warning us against the dangers of scientific utopianism. To this end, Huxley endows this society with features calculated to alienate his audience. He names the price of universal happiness as the destruction of our values, motherhood, home, freedom, love. To be fair, Huxley had me at his descriptions of the endless, socially mandated noise, designed to keep you from any sort of introspection. Likewise, Huxley ties biotechnology to fears of racism and eugenics. Using genetic engineering to make everyone healthy and well-adjusted is not the same as manufacturing a rigid hierarchy of genetically-preordained castes. In reality, there is no fundamental connection between soulless hedonistic existence and a scientifically engineered paradise. The lingering question of Huxley's legacy is if we

could have the best of both worlds, freedom, culture, family and love as well as the removal of all suffering.

We have learned to rationalize mental pain as good for us, developing character or increasing empathy. This is to remain sane in a world where there is no way to avoid it. This does not mean that mental pain could not be engineered away, or that it is morally wrong to do so. There's no causative link between Brave New World and the low social standing and funding of paradise engineering research. Huxley's views are more likely a symptom of wider societal disapproval than the cause. Still, "It's Brave New World" has become the thought-terminating cliché for arguments to improve society with anything from genetic engineering to designer drugs.

Is Huxley a fear mongering puritan denying us our lotus thrones, or a visionary saving us from a lever pressing hell? The answer, like most things, depends on your perspective. Someone struggling to get by on minimum wage, stuck in a house with abusive parents, or chronically lonely would jump at Huxley's Brave New World, where the drugs are free and the orgies continuous. Here even work is light and stress-free, calibrated to free you from the tedium of too much leisure. But Huxley was rich, well-educated, and a successful writer. His self insert, Helmholtz Watson, a handsome and gifted writer, is an "Alpha" genetically engineered to lead society. He is haunted by the feeling that society which has eradicated suffering has stifled his potential as a writer. He is foiled by Bernard, another "Alpha" whose short, low class stature leaves him a poor fit for the role society has assigned him. On Maslov's hierarchy of needs, Huxley is striving for self-actualization while the poor are fighting for their physical needs.

Unfortunately, Huxley was right about a lot of things. Despite declaring "science is everything", science is kept "chained and muzzled" by high-ranking politicians, to serve their own agenda. But lying, self-serving politicians are low-hanging fruit. Huxley describes how an equitable society made only of high-caste leaders dissolved, "all the people detailed for a spell of low grade work were perpetually intriguing for high-grade jobs, and all the people with high-grade jobs were counter-intriguing at all costs to stay where they were". (BNW 248) Now elite overproduction is fueling the rise of credentialism, where people burn years of their lives in universities, to scheme and backbite for tenure. Huxley envisions a future where "mother" is a bad word, and liberals are pushing to de-emphasize the family for "extended clan-networks". His ideas on genetic engineering and prolific drug use must be considered within the context of other, correct, predictions.

The logical extreme of paradise-engineering is not only eradicating people's suffering, but that of animals. Even the suffering of wild animals is not pre-ordained. Factory farming is a huge animal welfare issue, but interventions focus on reducing the need for meat via recruiting for vegetarianism, or improving physical conditions. There is next to no research on eliminating the pain and distress of domestic animals at the source, through wireheading or psychoactive pharmaceuticals. The sheer magnitude of human and animal suffering means any delay in a reduction of that suffering is a huge moral burden. To the extent to which *Brave New World* contributed to that delay, we must leave the suffering at Huxley's feet.

Enter Piper, the eponymous cat. He is orange, anxious and oblivious to the fact that his ass hangs out from beneath the kitchen chair. It's his favorite hiding place, ever since he was a kitten I could hold in the palm of my hand. Nowadays it takes both hands, but he still spends a lot of time there, wedging himself in whenever the neighbors stomp past the front door. Piper used to be a stray, and he's scared of nearly everything, including strangers, loud noises and new things. On the streets, it's a survival instinct. In my apartment, his fear vestigial, the appendix of the brain. Knowing that he'll never need it, should I drug him to remove his anxiety? Piper doesn't have a job, family responsibilities, or any moral duty to the society that made him. If it's moral to drug his anxiety away, isn't it also moral to give him heroin? If you love your cat, or your dog, shouldn't your goal be to drug them into eternal bliss? Would such an existence be worth living? What self actualization can a cat achieve?

The question is, with current technology, a moot point. Gabapentin, the anti-anxiety drug of choice for cats, is cleared through the kidneys, already the point of failure for a desert adapted animal. Yet when I consider it, more impactful than any list of side effects is the nauseated repulsion that comes from the image of Linda, pressured to drug herself to death on soma for the convenience of the society around her. As they justify to themselves, "every hour on soma is like a thousand years". (BNW 133)

As society moves towards total automation, more and more people are left without work, but also without self-actualization. Eventually, the drive to work and neurotic perfectionism of a model doctor or engineer become just as vestigial as darting beneath a chair at loud noises. This inspires the niggling suspicion humanity is in a race between automation and modifying ourselves to live in a completely automated society. The future teeters between engineering ourselves to be able to be content in paradise and societal collapse back to the hunter-gathers we're genetically primed to be. If we fall, I'm going to blame Huxley on the way down.

Huxley's work reflects the fears of the society he was steeped in. *Brave New World* centers drug use, female promiscuity, the war on religion and the destruction of the nuclear family. These fears have gone out of style, enough to make the modern, more politically correct reader uncomfortable. Still, the people who lambast Huxley as a eugenicist stopped reading when the Delta (low-caste) workers were described as "small, dark and hideous", (BNW 55) if they made it past the part where 70% of women were sterilized at birth. Certainly at twenty-five instead of fifteen I would have put down a book that declared "everybody (women) belongs to everybody else (men)". (BNW 46)

Brave New World's society is racist, or more appropriately, classist. But the characters shaped by this society are a stunning rebuke of the idea of genetic predestination. The low-caste workers were stunted with alcohol as embryos and conditioned since birth to accept society's place for them, but still have to be drugged into complacency with consumerism, entertainment and soma. Bernard, ordained to be an alpha, is an outcast due to his short stature and slim build. Even Helmholtz, the ostentatiously perfect alpha, is discontented.

Reading *Brave New World* as an adult, it's harder to defend. The plot meanders, the proposed society is unsettling and hits a little too close to home. Still, this was the cornerstone of the entire utopian genre, tangling with the idea of whether suffering eradicated through the perfection of society, and whether it should be. It is not good, but seminal. Seminal as the embryonic tissue of a genre, male centric and vaguely repulsive. Personally, as a teen *Brave New World* was the first book to change the way I thought. It helped me formulate questions I didn't know how to ask. The question at the core of Huxley's book is simple. Is the purpose of life to be happy? I still haven't figured that one out, but at least my cat can self-actualize.

Capital as Power: A Study of Order and Creorder by Jonathan Nitzan and Shimshon Bichler

Part A: Overview

1. A Need for Better Theory

If you are a well-educated person in the 21st century, you probably have conflicted views. On the one hand, the grand socialist project has had... problems... over the last century. Serious problems. Problems that kill and hurt people, and are really, really non-dismissable. For a bibliography, check [here](#).

On the other hand, the (main) alternative is capitalism. And that also sucks. A lot. If you haven't noticed this, you haven't been paying attention.

Capital as Power by Bichler and Nitzan does not even attempt to talk about an alternative system of government. However, it argues that a necessary precondition for radical system change is a new theory of economics. In their words:

Perhaps the key problem facing young people today is a lack of theoretical alternatives. A new social reality presupposes and implies a new social cosmology. To change the capitalist world, one first needs to re-conceive it; and that re-conception means new ways of thinking, new categories and new measurements.

The purpose of *Capital as Power* is to provide such a theoretical alternative. However such a thing is easier said than done. To start with, it is necessary to give a thorough examination of past attempts to put economics on sound theoretical feet.



Sound Theoretical Feet

Bichler and Nitzan provide some blistering polemic towards those who try to build the future without understanding this.

With some obvious exceptions, present-day leftists prefer to avoid 'the economy', and many are rather proud about it. To prioritize profit and accumulation, to theorize corporations and the stock market, to empirically research the gyrations of money and prices are all acts of narrow 'economism'.

To do these things is to fetishize the world, to conceal the cultural nuances of human consciousness, to prevent the critic from seeing the true political underpinnings of social affairs. Best to leave them to the dismal scientists. And, so, most self-respecting critics of capitalism remain happily ignorant of its 'economics', neoclassical as well as Marxist. They know little about the respective histories, questions and challenges of these theories, and they are oblivious to their triumphs, contradictions and failures. This innocence is certainly liberating. It allows critics to produce 'critical discourse' littered with cut-and-paste platitudes, ambiguities and often plain nonsense. Seldom do their 'critiques' tell us something important about the forces of contemporary capitalism, let alone about how these forces should be researched, understood and challenged.

Concordant with the spirit of this paragraph, Bichler and Nitzan devote much of the first half of the book to a critical dive into the history over the last two centuries of the "dismal science", both Marxist and Neoclassical.

I learned a *lot* about both neoliberalism and Marxism, and this part of the book would hold up as a good survey, even without the arguments they make for why the theories ultimately fail. It centers the analysis on an idea that both neoliberalism and Marxism are ultimately tied to *theories of value*. We are familiar with the neoliberal theory of value to such an extent that it is hard to even realize it is a theory. The neoliberal theory of value is that value comes from the utility that a good delivers to its consumer. Often this is how economics textbooks start, and they promise that the arguments that you can think of off the top of your head against this model have good counter-arguments, and in any case it's a useful model. Students with further questions are told that real economists use better theories than this, but they are too complicated to put in introductory textbooks. Bichler and Nitzan do a thorough job expounding on arguments that an intro econ student might think of, but could not come close to articulating in enough detail to make headway.

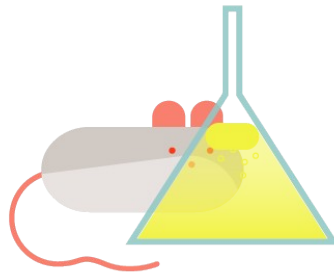
The Marxist theory of value is that value comes from the work that humans put into material goods. On the face of it, this makes a lot of sense. Ultimately, the limiting factor to production comes down to humans: no humans = no production. However, this theory also has holes in it.

The key to Bichler and Nitzans' arguments against both theories is that they cannot explain how capital accumulates, or provide a framework in which predictions about value can be made. So in short, the motivation for this book is that there are productive, empirical insights to be derived from a new economic point of view. I am used to alternatives to neoliberalism proposed for moral reasons, and it was refreshing to hear someone try to elucidate an alternative proposed for scientific reasons.

2. Business and Industry

One of the key parts of this new theory comes out of the theories of a historian named Thorstein Veblen (according to Wikipedia, Veblen coined the term "conspicuous consumption"). Veblen's big idea is that there is a fundamental distinction between

Business and Industry. Industry is the domain of the kind of people who build giant redstone contraptions in Minecraft or the kind of people who name their lab mice and talk to them in squeaky voices while cleaning up after an experiment that went on until 1am. It is both a collaborative activity and a competitive activity, but it is fundamentally built on creativity, curiosity and a desire to solve problems. Business is the domain of the kind of people who network at parties and care a lot about “corporate strategy”. The point of Business is profit and accumulation.



Trevor

Rather than theorizing capitalism as a perpetual struggle between classes, Veblen theorizes capitalism as a perpetual struggle between Business and Industry.

Modern capitalists are removed from production: they are absentee owners. Their ownership, says Veblen, doesn't contribute to industry; it merely controls it for profitable ends. And since the owners are absent from industry, the only way for them to exact their profit is by 'sabotaging' industry.

The *one thing* capitalism was supposed to be good at was high-quality goods at low prices, this is the promise dangled from the hands of every billboard in Times Square and every dense 800-page neoliberal economics book. But actually, business subverts the production of quality goods at low prices for the purpose of profit. One of the most obvious examples of this is intellectual property; industry is siloed into many companies who cannot freely remix and use each other's designs. Open source software, even though it is massively underfunded compared to proprietary software, often manages to punch above its weight because of the superior development model of sharing, and because it is not sabotaged by profit considerations.

If nothing else, *Capital as Power* is worth reading for the wealth of examples of this conflict.

This is perhaps the reason why early in the twentieth century the automobile companies bought and dismantled 100 electric railway systems in 45 US cities (Barnet 1980: Ch. 2). And it is also why these companies have long shunned any radical change in energy sources. The electric car, first invented in the 1830s, predates its gasoline and diesel counterparts by half a century, and for a while was more popular than both (Wakefield 1994). But by the early

twentieth century, having proved less profitable than the gas guzzlers, it fell out of favour and was forcefully erased from the collective memory. Then came intolerable pollution, which in the 1990s led the state of California to mandate a gradual transition of automobiles to alternative energy. Complying with the new regulations, General Motors had its engineers quickly develop a highly efficient electric car, the EV1. But fearing that this gem of a car would undermine profit from their gas guzzlers, the company's owners, along with owners of other concerned corporations in the automotive and oil business, also invested in an orchestrated attempt to defeat the California bill. When the regulation was finally overturned, every specimen of the EV1 was recalled and literally shredded (Paine 2006).

The idea of a distinction between Business and Industry was not the most provocative idea in *Capital as Power*, nor the one with the farthest-reaching implications, but it was the one that stuck in my head the most. It seems to me to be a very productive way of thinking, and sums up a lot of stuff I didn't have words to describe before. But to Bichler and Nitzan, this is merely a springboard for a much larger theory.

3. Accumulation of Power

Here, Bichler and Nitzan follow the ideas of a historian named Lewis Mumford. Mumford takes us back all the way to the beginning of what we now call civilization, in the Nile River delta. His conceit is that the first "technology" was not mechanical or chemical, it was social. The organizational structure of ancient Egypt, with its intricate hierarchy of politics and religion meshed together, was a form of power previously unmatched in its ability to change its surroundings and to persist through time.

And here I add my own analysis. It was impressed on me thoroughly in ninth grade ancient history by my most excellent history teacher Audrey Budding that the one of most common threads through history and one of the most important questions to ask about a society is how legitimacy of the ruling class is achieved. In this lens, the novel technology of Egypt was the ability to give its rulers, through religion, tradition, force, and bureaucracy, enough legitimacy that they could impress their will on a massive population.



The Original Mega-Machine

Mumford calls this new technology the “mega-machine”, and Bichler and Nitzan take an interesting romp through some of the mega-machine’s greatest hits throughout the years since the fall of Egypt.

According to them, the most recent incarnation of the mega-machine is the entirely novel *quantification of power through capital*. With the dazzling mathematics of the market, the mega-machine has reached heights of sophistication that the gold-plated pharaohs of yester-millennium could only dream of. The mechanism by which the ruling class exerts power is tightly woven into the daily fabric of the lives of their thralls, and legitimized by every interaction that becomes a transaction.

At this point, you are probably thinking “Yes! This is the hot tea that I signed up for when I clicked on a link that claimed to lead to a book review of a book called *Capital as Power!*”

I hate to disappoint you, but this is the end of the introduction, and we’re not going to get back to the real juicy stuff until after about a thousand or so words droning on about theories of value and dead white men.

You see, I’m trying to give you a picture of what it is like to read this book, and the experience of having a tantalizing insight dangled in front of you but then being forced to read far more history and statistics than you would really like to understand it is essentially all of *Capital as Power*.

Summary of Part A

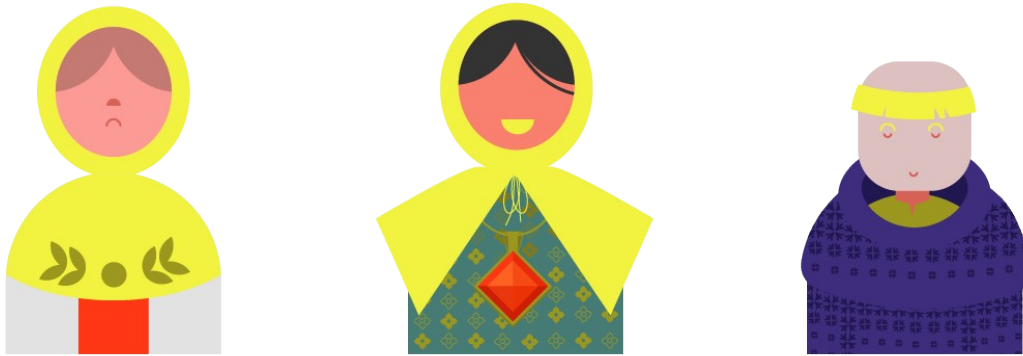
- Capitalism is bad, but before we can improve it, we need to understand it scientifically.
- Current economics has some deep flaws in this regard.
- Cultural critiques are dumb.
- Business and Industry are two distinct things.

- Capitalism is, like, Egyptian pharaohs but with more numbers.

Part B: Dilemmas of the Dismal Science

1. Politics and Economy

Before the industrial revolution, one could make a decent argument that political power and economic power could be separated. One strong point in favor was the separation of nobles, merchants, and clergy. Certainly, nobles often happened to be rich, but their wealth mostly derived from land-ownership, and their political power was mainly derived from their birth and connections, rather than their wealth (though, connections and wealth could go hand and hand). Merchants, even though they could have influence through their wealth, were excluded from positions of political power. And the clergy, in some cases more important than the nobles or the merchants, were also quite separated from both.



Noble, Merchant, Clergy

It is from this situation that the “original sin” of economics derives: the belief that economics is productively analyzed outside of political context. The original liberals relied on this duality to expound on their idea that the ideal government simply facilitates the movements of the market.

Marx was the first to question this belief, claiming that it was only through political oppression that industrialists were able to achieve their economic exploitation. However, he still makes an essential distinction between the two, and the Marxist viewpoint is that the contradictions between the political and economic spheres are what will eventually bring down capitalism and usher in communism, where politics and economics will finally be united.

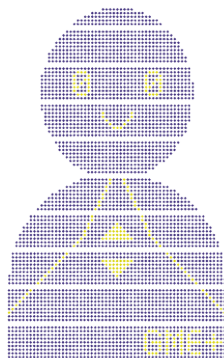
Bichler and Nitzan take as a jumping-off point this faulty duality, and use it to explain how various problems come up in both theories. For instance, the pure “Newtonian” laws of microeconomics were eventually forced to be revised by the Great Depression, and the new science of macroeconomics was introduced to account for this. Since then, the systematic differences of the real world from the “spontaneous equilibrium” of the

market have been accounted for by an ever growing pageant of distortions, applied ad-hoc. But just as Ptolemaic astronomy eventually drowned under the weight of its many epicycles, so is neoclassical economics struggling under the weight of all of the various compensations needed to account for politics in a theory that assumes politics is out of the picture.

In Marxism, a similar problem arose as the competition-rich environment where Marx originally made his theories gave way to a new monopolistic capitalism. Without competition, the tendency of prices of goods to correlate with the prices of labor, and for profits to equalize among different sectors no longer held true. To account for this, Neo-Marxists developed a new theory that attempted to bring power back into the picture. However, like the man who cannot light his cigarette without putting down his teacup, in order to do so the labor theory of value had to be jettisoned, and Neo-Marxism became unmoored from the theoretical framework that birthed it. When the stagflation of the 70s and 80s hit along with the breakup of “state capitalism”, Neo-Marxism ceased to be an accurate description of the world, and leftists attempted to move back to original Marxism, decrying the period of monopoly as a “historical blip”. However, there was little left of the original theory that still made sense, and many Marxists moved to cultural critiques, abandoning the original attempt of Marxism to put the study of capitalism on a scientific basis.

2. The Nominal and The Material

Another issue that plagues both Marxist and neoliberal analyses is an attempt to make a distinction between what I might call “the map” and the “territory.” On the Marxist side, this comes up as a distinction between “fictional” and “real” capital. “Real” capital is owned by industrial capitalists, who employ productive labor to create surplus value. I imagine that Marx was thinking of factories here. “Fictional” capital, on the other hand, is capital owned by commercial and financial capitalists, who merely appropriate the value generated by productive labor. Intuitively, this is a theory that one is inclined to be sympathetic to. On the one hand, you have farmers and laborers, who are clearly producing surplus, and on the other hand you have stock brokers, who clearly are just siphoning off the top. But to actually give a sturdy definition of what is “productive” and what is “unproductive” is not at all easy.



On the other hand, the liberals face a similar dilemma as they argue that the market ultimately represents the movements of actual industrial processes. Bichler and Nitzan argue that liberals were originally motivated to do this in order to make the argument that, unlike nobles who acquired their wealth by looting or birth, the new rich acquired their money by work. If ultimately the market is just a representation of real processes that bring material changes in the world, then business deals are as legitimate a form of work as manufacturing and farming, and so just as the laborer is entitled to the work brought in by the sweat of his brow, so is the businessman entitled to whatever rewards he can manage to make by fair participation in the market.

However, this has its own problems. The movements of capital markets are empirically quite difficult to correlate with the movements of the underlying material assets that those markets are supposed to represent. Again, neoliberals employ a pageant of epicycles to explain away this, but there are only so many epicycles that a scientist should accept.

In short, applying the “map and territory” analogy to money and “real” assets is bad for two reasons. The first is that it is very hard to draw a line between the two, and the second is that even when we can tell that something is a map, it’s very difficult to figure out what it is a map of.

3. What is Value?

The last problem is more philosophical. It’s best introduced in Bichler and Nitzan’s own words.

To study the rationalist order of capitalism without quantities is like studying feudalism without religion, or physics without mathematics. According to Marx, and here he was right on the mark, capitalism, by its very nature, seeks to turn quality into quantity, to objectify and reify social relations as if they were natural and unassailable. In this sense, a qualitative theory of value necessarily implies a quantitative theory of value; it means a society not only obsessed with numbers, but actually shaped and organized by numbers. This organization is the architecture of capitalist power. To understand capitalism therefore is to decipher the link between quality and quantity, to reduce the multifaceted nature of social power to the universal appearance of capital accumulation. The two aspects of the theory rise and fall together. If one is proven wrong, so is the other.

With this passage, Bichler and Nitzan set a high bar for theories. It is not enough for a theory of capitalism to give a qualitative account of value because the nature of capitalism is quantitative. It is on this cross of numerics that the Marxist “abstract labor” and the neoclassicist “util” are ultimately crucified.

They lay the problem out in the following way. It is well understood that wealth is not well-measured by current market value, because of inflation. Economists get around this by coming up with a “price index”, which measures how much a “standard bundle” of goods and services would cost at different times. By this mechanism, comparisons of value across time can be measured. However, the definition of such a “standard bundle”

is highly problematic. If this standard bundle were, say, a single transistor, then an economist would conclude that we are over a million times richer than we were in the 1960s. This is clearly not true. Of course, an economist would not calculate a price index in this way, but almost any commodity has similar problems, though less extreme, from roast passenger pigeon to steel.



A Standard Bundle of Goods

In practice, some reasonable judgement is made on which commodities to include. But on what basis is this judgement made? The theoretical underpinnings of neoclassicism are that things are ultimately denominated in “utils”. This explains why we don’t simply multiply by the number of transistors: utils don’t scale linearly with the number of transistors. This explains why roast passenger pigeon would go for a high price now; it’s scarcity would cause someone to derive great utility from their signaling of wealth by eating it, or great utility from their curiosity in how it tastes.

As a qualitative theory, utils certainly have a lot going for them. I buy groceries because I derive utility from eating them. I pay rent because I derive utility from having a roof over my head. But even the most ardent proponents of the new utilitarianism balked at doing actual calculations.

In fact, they admitted quite openly that universal utility is impossible to measure and, indeed, difficult to even fathom. The interesting thing, though, is that this recognition did not deter them in the least. ‘If you cannot measure, measure anyhow,’ complained their in-house critic Frank Knight

Instead, neoliberal economists developed a theory of “revealed preferences.” This theory says the following:

‘Utility is the quality in commodities that makes individuals want to buy them, and the fact that individuals want to buy commodities shows that they have utility’ (Robinson 1962: 48).

We wanted to compare wealth over time. Prices change due to inflation, technological improvements, and a whole host of other things, so we need a theory of value that allows us to make this comparison. Utility provides that theory of value. How do we measure utility? Prices. Whoops...

I want to make it explicit that Bichler and Nitzan aren't making the argument that utilitarianism is not a good basis for morality. This kind of argument would not be novel or relevant. They are making the argument that utilitarianism is a bad basis for a theory of capitalism because it fails to make quantitative testable predictions.

While I'm clarifying things, I might as well reiterate (though it should go without saying) that anyone wishing to refute any of the arguments that I set out here, should not take my presentation as canonical in any way; Bichler and Nitzan's arguments are far more comprehensive than my brief summary. Although I promised that I would bore you with history and statistics, I really don't have the space or time.

A similar circularity happens with the labor theory of value. I won't go into as much detail, because I think it's somewhat widely accepted that the labor theory of value doesn't really work, but I encourage the interested reader to read the relevant parts in "Capital as Power"; I learned a lot about what the labor theory of value actually means that I had never really had explained to me in such detail.

Summary of Part B

- Separating politics and economy leads to contradictions
- Separating out "real" from "fictional" capital also leads to contradictions
- Assuming that prices are caused by some sort of externally-defined value leads to contradictions.

Part C: The Machine

Moloch whose blood is running money!

1. Internal and External Logic

If we cannot use utility or abstract labor, then by what means can we determine value? Bichler and Nitzan do not have an answer to this question. They instead say that it is the wrong question to ask. It makes no sense to try to anchor the fluctuating numbers of human society to some fixed and eternal quantity. Rather, prices are just another expression of society.

According to Cornelius Castoriadis (1984), this alternative was articulated some 2,500 years ago, by Aristotle. Equivalence in exchange, Aristotle argued, came not from anything intrinsic to commodities, but from what the Greek called the *nomos*. It was rooted not in the material sphere of consumption and production, but in the broader social–legal–historical institutions of society. It was not an objective substance, but a human creation.

This articulates the idea that capitalism exists on a continuum. The idea that a gift economy or dictatorship is based on societal context is totally uncontroversial; why should the theory of a transactional economy be any different?

Consider the ratio between the price of petroleum and the wages of oil rig workers; between the value of Enron's assets and the salaries of accountants;

between General Electric's rate of profit and the price of jet engines; between Halliburton's earnings and the cost of 're-building' Iraq; between Viacom's taxes and advertisement rates; between the market capitalization of sub-prime lenders and government bailouts. Why insist that these ratios are somehow determined by — or deviate from — relative utility or relative abstract labour time? Why anchor the logic of capitalism in quanta that cannot be shown to exist, and that no one — not even those who need to know them in order to set prices — has the slightest idea what they are? Isn't it possible that these capitalist ratios are simply the outcome of social struggles and cooperation?

The title of this section comes from my personal spin on this. Euclid tried to define all of the geometrical terms that he worked with rigorously. However, such an attempt was inevitably circular. Modern mathematics affirms that the proper way to axiomatize math is to start with undefined terms, which derive meaning through their relation to each other. That is, ultimately mathematics cannot be given an "external" grounding; it must be defined through "internal" means.

This is also well-understood in the linguistic sphere. An objective definition of a word does not exist; all that exists is the totality of that word's relationship with actions and with other words. Analogously, it makes sense that prices and capital and markets are not objectively defined quantities, but exist only in relation to the larger context of civilization.

This resolves the dilemmas of part B. Politics and economics are clearly not separated; the study of one is the study of the other. The private ownership of "financial" assets is just as "real" as the private ownership of material goods; they are simply relations that exist in a certain societal context. And finally, there is no objective standard of value.

But wait, you are surely thinking, isn't this giving up? After crucifying neoliberal and Marxist theories on a cross of numerics, it's a pretty poor showing to have the theory that's supposed to replace them be all "societally-determined" woo.

The answer is that productive empirical tests and theories can be produced from this mindset. Abandoning an "external logic" of capitalism does not mean that scientific inquiry has to pack up and go home. Rather, it frees us to look for theories that are embedded in specific contexts and divorced from pretensions of universality.

An analogy can be drawn here to machine translation: the "rules-based" systems of the early AI days are no match for the deep learning systems of today, and the transition was predicated on an understanding that the meaning of a word can't be pinned down with formal rules (and of course also predicated on access to GPUs).

Additionally, this mindset allows us to salvage a great deal of neoliberal and Marxist economics. When properly contextualized and removed from their moral and theoretical underpinnings, these theories can have great empirical success.

The famous physicist and captain of the high seas Robert Hooke believed that the law that he discovered governing the motion of a spring, $\ddot{x} = -kx$, where x is the displacement from rest position, was a new law of nature. However, it actually turned

out to just be a good approximation in certain cases, and we still use Hooke's law in engineering because of this

I think that many economists are all too aware that they are just using an approximation. And if I wanted to know the answer to a given question, I would look to expert economic judgement. This is because a highly tuned misguided theory often gives much better results than a poorly tuned well-founded theory. I recently bought a subscription to the Economist, and despite its heavy neoliberal bias, I still think it gives a clearer picture of the world than many other news sources; a large helping of meritocracy goes a long way. However, the way that human brains work is that understanding is layered. After working with principles and laws for a long time, we forget what the underlying assumptions that lead us to come up with those principles and laws were, and are therefore less able to ascertain when those principles and laws no longer hold. Our intuitions about prices are developed by going to the grocery store, not by buying companies, and we should not trust principles just because they seem to make sense in the micro-scale.

In this way, implicit assumptions about the nature of economics radically change the type of models that are even considered, and also radically change political views. Therefore, although this indictment of neoliberal theory does not in my mind invalidate expert scientific consensus in many areas, it does undermine *political* arguments based on neoliberal/Marxist theory, and it opens the door to new scientific ideas. And both of these are needed if this new theory is to be a success; to make a dent in neoliberal consensus where "cultural" theories have failed it must open up previously poorly understood areas to empirical analysis, and to better guide society it must have political implications based on that new empirical analysis.

2. Capital and Capitalization

In the previous section, we moved a bit far afield from the review. We now move back to the book as it attempts to examine the "internal logic" of capitalism; the processes and beliefs that keep it afloat.

Bichler and Nitzan argue that the central process of capitalism is, fittingly, *capitalization*. Capitalization is the process of taking an asset that is expected to produce a certain stream of future profits and assigning a current price to it. This is achieved through a *discount rate*, where future profits are valued lower than present profits. That is, we assume that 1 dollar now is worth the same as r dollars in a year. Mathematically, we can then say that if the profit from an asset n years in the future is p_n , then the value of that asset now is $\sum_{n=0}^{\infty} r^n p_n$. If $p = p_n$ is constant, then this reduces to

$$\sum_{n=0}^{\infty} r^n p = \frac{p}{1-r}$$

This formula is called the "discounting/capitalization formula". This idea of a discount rate (and implicitly this formula, and variations of it), have been in use since the fourteenth century, where it was first introduced by Italian merchants (according to

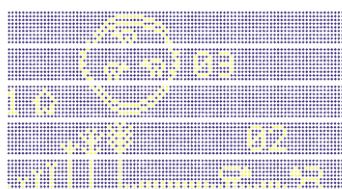
Bichler and Nitzan). As the centuries went on, it spread farther and farther, although not everyone really did the math “correctly”.

In 1907, Irving Fisher proposed that this discounting logic was in fact universal.

It is evident that not bonds and notes alone, but all securities, imply in their price and their expected returns a rate of interest. There is thus an implicit rate of interest in stocks as well as in bonds.... It is, to be sure, often difficult to work out this rate definitely, on account of the elusive element of chance; but it has an existence in all capital.... It is not because the orchard is worth \$20,000 that the annual crop will be worth \$1000, but it is because the annual crop is worth \$1000 that the orchard will be worth \$20,000. The \$20,000 is the discounted value of the expected income of \$1000 per annum; and in the process of discounting, a rate of interest of 5 per cent. is implied.

Bichler and Nitzan are really good at picking good quotes from old economists, so I can't resist giving you another one:

The primitive economy in its choice of enjoyable goods of different epochs of maturity, in its wars for the possession of hunting grounds and pastures, in its slow accumulation of a store of valuable durable tools, weapons, houses, boats, ornaments, flocks and herds, first appropriated from nature, and then carefully guarded and added to by patient effort — in all this and in much else the primitive economy, even though it were quite patriarchal and communistic, without money, without formal trade, without definite arithmetic calculations, was nevertheless capitalizing, and therefore embodying in its economic environment a rate of premium and discount as between present and future. (Fetter 1914b: 77)



Livestocks

After this quote, Bichler and Nitzan quip

In short, if human beings were indeed made in the image of God, the Almighty must have been a bond trader.

However, despite these enthusiastic embraces of capitalization from the dismal science, the general public were not convinced, and the capitalization formula was not yet embraced. It was not until the dazzling onslaught of complexity in corporate finance that started to unfold in the 1930s and 1940s that the capitalization formula was firmly grasped as a principle with which to evaluate and price assets and make sense of the growing chaos. Later on, risk was incorporated formally into the model, and these

modern practices of corporate finance became so firmly entrenched that one would be entirely forgiven if one supposed that they were laws of nature.

For me, this was an interestingly different way of looking at capitalism. Capitalism is typically presented as a system that embraces markets, whereas this presents capitalism as a logic for how to price things in the context of markets. In theory, another political system could use markets without using the same pricing logic.

Bichler and Nitzan then devote a whole chapter to showing that capital within the modern capitalist system is determined wholly by variants of the discounting formula, and that capital is not necessarily correlated with any sort of “real capital” in terms of physical assets. Unfortunately we don’t have time to get into this, but this is where they start pulling out the graphs and data, and start to show that the assumptions that economics textbooks make in correlating capital with any sort of measurable, tangible asset are empirically unfounded.

3. Profit and Differential Profit

If capital is “capitalized future profits”, then what are profits? Where do they come from, and how do they relate to accumulation of capital?

Recall the framework of “Business and Industry” that Veblen came up with. According to Veblen, when Industry is allowed to run unchecked, there are no profits. Competition makes the price and cost of a good equalize exactly. In order to make profit, a firm must somehow restrict industry, or “sabotage” it, in the case of Veblen. For instance, in section A.2 we saw that the profits of automobile companies directly correspond to their ability to sabotage alternate forms of transportation. To sum up, capital measures and “discounts” a firm’s ability to sabotage industry.

This goes along with a theory of property rights that sees property rights as fundamentally exclusionary. My ownership of a house does not enable me to use it, it only serves to disallow you from using it. Capital measures the access of a firm to revenue streams *that other firms cannot access*. A textbook of 21st century materials engineering might be worth billions to a firm in the 1960s that could control access to its knowledge, but once it becomes common knowledge, it cannot be counted as capital even though the firm’s ability to access it has not changed at all.

As usual, Bichler and Nitzan write about this in a very elegant way.

Business, like other power institutions throughout history, can force people to act, but it cannot make them productive. Moreover, productivity as such, being socially holographic and therefore open and unrestricted, cannot generate a profit. The only way for capitalists to profit from productivity is by subjugating and limiting it. And since business earnings hinge on strategic sabotage, their capitalization represents nothing but incapacitation. In this particular sense, capital, by its very construction, is a negative industrial magnitude.

Even human and relationship capital can be viewed in this lens. When human capital is weighed on the balance sheet, what is being measured is the firm’s ability to divert the

output of those brains away from increasing the common knowledge and towards increasing the profit of the firm.

However, as power is inherently relative, the way it accumulates is not by profit, because a rising tide that lifts up all boats does not change their relative standing. Instead, capital accumulates by *differential* profit. In other words, firms are unconcerned with outrunning the bear market, they are just concerned with outrunning you.

Bichler and Nitzan phrase this culturally, talking about the drive to beat the average among stockbrokers, but one could easily think about this evolutionarily. The firms that have managed to keep a consistent rate of growth that is above the average have, by the miracle of exponential growth, become those firms that dominate the economy.

In short, differential profit measures the ability of a firm to sabotage industry *more than everyone else*. This race to beat the average is what determines which firms end up on top, and which CEOs can buy the most expensive yachts. Not efficiency, not productivity, not innovation, but the ability to sabotage industry more than everyone else.

Summary of Part C

- It makes more sense to think of prices as numbers which have internal, not external significance.
- The central process of capitalism is capitalization of future profits (measured in prices)
- Profit measures the ability of a firm to sabotage industry
- Firms grow when they can sabotage industry more than other firms
- Capital measures the ability of a firm to sabotage industry, discounted over time.

Part D: The Machine in Context

1. Economic Implications

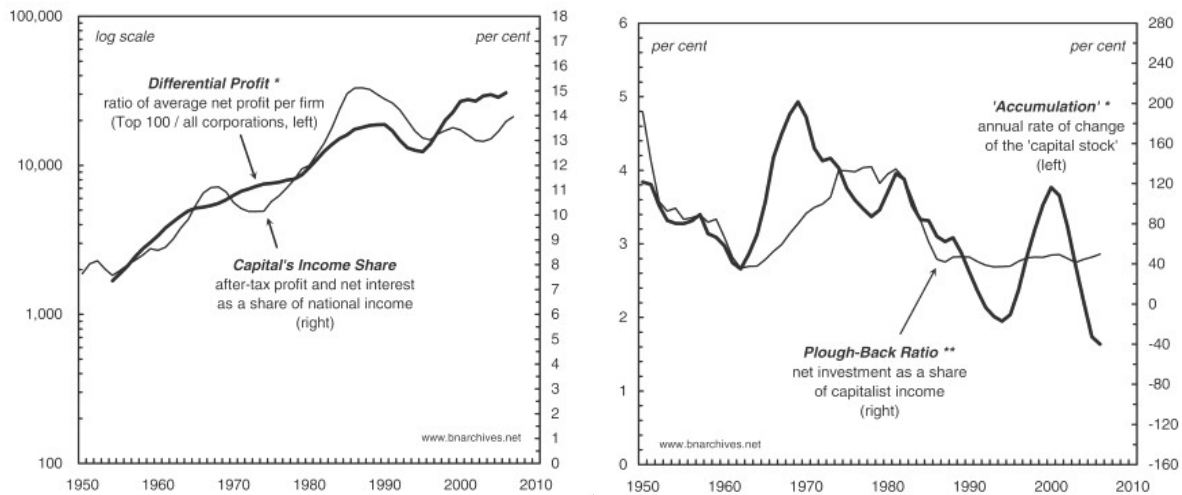
In this section, we talk about some of the predictions that this theory makes, and how well they align with reality.

First of all, no review of Capital as Power would be complete without mentioning their discussion of what they call “Dominant Capital.” Dominant Capital is what they call the conglomerate of the very largest companies in capitalism, the companies who have managed to squeeze above-average profits out of industry year after year and have an outsized influence on not only the markets, but also the political process.

As a rough approximation of dominant capital, Bichler and Nitzan consider the top 100 companies as measured by capitalization. Using this, Bichler and Nitzan shows that while various other measures of growth have not had clear trendlines, accumulation of capital by the top 100 companies has been happening at a steady rate since WWII.

This growth is not driven by expansion of industry. Bichler and Nitzan argue that expansion of industry means loss of control by capitalists, and thus is uncorrelated to accumulation of capital.

To back these arguments up, Bichler and Nitzan make lots of graphs where two trends on different axes are lined up.



These sort of graphs are actually pervasive throughout *Capital as Power*, and although they were suggestive, I'm kind of skeptical of arguments that pull out variables from thin air and show that they are correlated. Presumably Bichler and Nitzan had many different datasets that they could work with; it's not too hard to find trends that are correlated if you squint. And additionally, as always causation is hard to tease out; the cycles/trends in these graphs could be coming from common causes that have nothing to do with this theory.

As a current statistics grad student, I was both pleased and disappointed at the general level of math in this book. That is, there is very little math. From my perspective, this is *far* better than coming up with a lot of ad-hoc models with lots of assumptions; better to have a tight qualitative analysis than a sloppy quantitative one. On the other hand, I was kind of hoping that there would be something a little more precise than "these graphs kind of look similar." I haven't read much of the rest of Bichler and Nitzan's work though, so perhaps they develop these theories more rigorously elsewhere.

Another consequence of the frame of *Capital as Power* is a new look on inflation. To start this analysis, Bichler and Nitzan claim that one of the most common modes of power is *price setting*. In neoliberal economics, firms are typically portrayed as having to take prices dictated by the market, but in fact in order to achieve a "normal rate of return", firms must exert power and set prices higher than what a truly competitive market would bear.

As a consequence of the former point, inflation is the result of power struggles between businesses, and is fundamentally redistributionary, rather than being an expression of a

growing economy. This explains stagflation; price wars without industry growth makes a lot of sense; business expresses power by both restricting industry and raising prices.

Additionally, Bichler and Nitzan claim that in the US for the last 50 years, dominant capital accumulates during periods of inflation, where dominant capital can raise prices faster than everyone else, and slows accumulation during periods without inflation. However, this is not necessarily true in other countries, in which different redistribution patterns happen during inflation. That is, the US economy has become a good engine for the accumulation of dominant capital, but this is by no means universal.

The final consequence that I will mention briefly is *mergers and acquisitions*. Apparently much of the growth of dominant capital in the last few decades has been mediated by mergers and acquisitions, not green-field growth. This makes sense because mergers and acquisitions allow for more concentrated power and more control over industry, and green-field growth has the potential for letting industry run away from business.

In general, I would be very interested for someone with a better understanding of economic indicators than I to take a look at the later chapters in *Capital as Power* and tell me whether or not Bichler and Nitzan are cherry picking their data or not, and whether their analysis is accurate; this was the part where they really started to lose me, but also the part where all of their claims about the empirical verifiability of their theories rest.

For this book review, however, we must move on!

2. The Space of Political Systems

Capital as Power has a subtitle which I have not mentioned yet. Its full title is “*Capital as Power: A Study of Order and Creorder*.” What is creorder? I’ll let the authors define this for you

Historical society is a creorder. At every passing moment, it is both Parmenidean and Heraclitean: a state in process, a construct reconstructed, a form transformed. To have a history is to create order — a verb and a noun whose fusion yields the verb-noun creorder.

A creorder can be hierarchical as in dictatorship or tight bureaucracy, horizontal as in direct democracy, or something in between. Its pace of change can be imperceptibly slow — as it was in many ancient tyrannies — yielding the impression of complete stability; or it can be so fast as to undermine any semblance of structure, as it often is in capitalism. Its transformative pattern can be continuous or discrete, uniform or erratic, singular or multifaceted. But whatever its particular properties, it is always a paradoxical duality — a dynamic creation of a static order....

The use of this idea is to situate capitalism within a broader space of creorders. Capitalism is characterized by the quantification of power through a market, and the accumulation of power through exclusive rights. The creordering of the market, and of these property rights, however, is accomplished through the very powers that the markets and property rights affirm. That makes the exact power structure of a capitalist

society very fluid. The internal logic of one capitalist society are due to these power relations, and though the power relations may be mediated through the market; there are not “economic laws” that force the logic of one capitalist society to be the same as that of another capitalist society. We have mainly been analyzing the logic of the U.S. economy over the last 50 years, and Bichler and Nitzan are able to separate out the universal framework from particular features of this system.

One example of using “particular theories: is Bichler and Nitzan’s idea of the “petrodollar-weapondollar coalition” vs. the “technodollar-mergerdollar alliance.” They posit that during the Cold War, dominant capital mainly centered around oil extraction and petroleum-dependent industry, and additionally weapons manufacture: the “petrodollar-weapondollar coalition”. These two sectors were dependent on each other because oil required certain international relations that the defense industry was happy to supply. However, as the Cold War wound down in the 90s, dominant capital shifted towards technology and consolidation of power through mergers, which they call the “technodollar-mergerdollar alliance”. This new coalition seemed to observers to herald a new economy centered around growth through high-tech knowledge and industry.

However, with the dotcom crash and the new wars in the Middle East, dominant capital shifted back to “petrodollar-weapondollar”. Depending on where dominant capital was centered, different logic applied.

Bichler and Nitzan developed this theory of “petrodollar-weapondollar” and “technodollar-mergerdollar” back in the 90s, and claim that they were able to successfully predict changes in trends based on this framework that other forecasters were unable to see at the time, in published articles. This should be easy enough to verify for anyone who wants to dig through their old papers, and seems like a strong indicator that they know what they are talking about, at least in this area.

The point that I am trying to make is that, unlike the laws of economics which mostly claim to be universal across time, the strength of Capital as Power is that they can identify what things are true about some periods, and not of others, and integrate these assumptions into their models. In other words, rather than being a general theory of economics, Capital as Power is a general theory of the space of possible capitalist politics, or as Bichler and Nitzan seem to be so happy to coin, a general theory of possible capitalist creorders.

3. Connections with Other Theories

Reading this book, I had the overwhelming desire to introduce Bichler and Nitzan to James C. Scott, the author of “Seeing Like a State.” There seems to be a very strong similarity to how each theorize power. That is, power mainly serves to organize society in such a way that its resources are more easily extracted. However, Bichler and Nitzan are not anthropologists, and James C. Scott is not an economist, and as such each of their analyses is limited by domain.

Specifically, I think that Capital as Power could be greatly improved by a discussion of “legibility.” It seems like one important asset that is discounted implicitly by capitalization

is the “legibility” of industry to a business. If an industry is illegible, it is much more difficult to extract profits.

Conversely, James C. Scott is heavy on examples and creates good language to describe situations that were previously harder to describe, however he lacks empirical/numerical theories built off his framework. Using the framework of Capital as Power, it could be more possible to somehow correlate legibility of a certain domain to the stock price of businesses with interest in that domain.

Finally, the idea that I have had in my head ever since reading Seeing Like a State (or more accurately, the idea that I have had in my head ever since reading Scott Alexander’s review of Seeing Like a State) was that there should be some way of talking about all of this using some variant of statistical mechanics. If power is the ability to create order, and order is the absence of entropy, then we should be able to talk about power and creorder in the framework of a stochastic dynamical system, of which we can measure the entropy at certain points of time.

The second law of thermodynamics says that entropy always increases in a closed system. However, in an open system, such as the Earth which is constantly receiving sunlight and radiating out waste heat, this does not have to hold. And in fact, all of evolution and human society is proof that the second law of thermodynamics does not in fact hold in an open system.

This semester, I’m finally taking stochastic mechanics (up to now I’ve only had a vague idea of what it actually is), so maybe I’ll be able to say more after I know what I’m talking about. But if you, dear reader, are catching a glimpse of this vision that I have, and you know something about statistical mechanics, I ask that you keep this idea in your head and toy with it over the next decade. A mathematical framework for creorder would dramatically expand the set of questions that scientists can research about the world.

Summary of Part D

- The Capital as Power theory of the capitalist machine has many concrete explanations of graphs of economic indicators over the last 50 years in the United States. Whether or not these are just-so stories must be left to a more-informed reader than I.
- Capital as Power provides a framework for talking about specific internal logics of capitalism, and relating them to each other.
- Capital as Power seems to be hitting at a much larger theory, which is also talked about by James C. Scott, and should be mathematically modeled by stochastic mechanics.

Should You Read This Book?

Capital as Power is long, but extremely full of content. There are some large points in the book that I didn’t talk about at all, and that which I did talk about was highly condensed. After reading both Seeing Like a State and the slatestarcodex review of Seeing Like a State, I thought that Scott Alexander managed to capture fully the main

points, but I would definitely say that this review has not captured in full the main points of Capital as Power. Put in a more positive way; this review has not spoiled Capital as Power for you!

Additionally, Bichler and Nitzan write in a very engaging way; not necessarily easy to read but certainly action-packed. And there are many, many interesting historical nuggets in the book, like the history of GM's EV1 car that I referenced in the introduction.

And finally, although Capital as Power is long, there is a "middle way" (between reading this review, and reading the entire book). The first chapter, which is around 80 pages on my ereader, contains a summary and overview of the whole thing, and Bichler and Nitzan have a [free copy](#) available on their webpage, so you don't even have to feel bad about buying an entire book. So go do that!

Che Guevara: A Revolutionary Life by Jon Lee Anderson

The book makes one think about the role of personalities in history: Cuba became communist just because of the Argentine Che Guevara. He planted the idea of communism in the mind of Fidel Castro. Fidel wasn't a communist in the beginning. Fidel was an ambitious, Che an idealist. Fidel wanted power, Che wanted a better world. Fidel had bourgeois leanings: he liked comfort, good food and beautiful women. Che pushed correctness and self-denial close to asceticism. And still, Che, *le chevalier sans peur et sans reproche*¹¹ of communism, had a far bigger negative impact. Without Che, Fidel would have been just an ordinary Latin American dictator.

Don't think that this book only shows Che Guevara's qualities, although some people consider it a hagiography. On the other hand, I've seen Marxist sites that consider Jon Lee Anderson as a detractor who tainted Che Guevara's memory. Because Anderson tries to be objective: he also writes about the shootings ordered by Che Guevara and about his disastrous impact for the Cuban economy, as a Minister of Industry (He was more communist than Moscow, he didn't accept even the mild financial autonomy of the companies and financial rewards for workers allowed in other communist countries, he thought that the economy could work only on moral motivations.) and about the toughness of this symbol of rebellion with his subordinates (for serious faults, he punished his fighters with days without any food, and his subordinates at the Ministry, with months in the labor camp Guanahacabibes). Although the author lived for years in Cuba, when researching the book, the book is far from being pro-Cuban propaganda. I won't be surprised to find that it's banned in Cuba – actually, I would be surprised if it isn't.

Of course, Anderson isn't totally objective. An article about Che Guevara's victims accused the book of having more pages about Che's first love than about the trials and executions of those accused of involvement in dictator Batista's repressions, organized by Che Guevara in the prison-fortress La Cabaña in the first months after the victory. I noticed that in the La Cabaña part, Anderson only quotes sources on Che's side, subordinates involved themselves in the revolutionary tribunals. There were 55 executions in 5 or 6 months. To put things into perspective, the author says that, about the same time, Raul Castro (Fidel's brother and later successor) executed 70 soldiers in

11 *le chevalier sans peur et sans reproche* - the knight without fear or blame (French)

just 1 day. There were also executions during the fights in the mountains: informers, deserters, robbers, rapists – here the source is Che himself, who wrote everything in his diary, including how he shot a traitor with his own hand, where the bullet entered and where it exited the skull (using anatomical terminology – he was a doctor). To think that he chose the path that led to all that from compassion, moved by the suffering of the poor...

To his own conscience and to his old friends shocked by his metamorphosis in a “killing machine” (like Alberto, the guy with the motorcycle), he justified his actions on utilitarian grounds, speaking about the future. If you don’t kill them, they kill you. Yes, they also have wives and children, but you have the chance to save thousands of future children from poverty and hunger. As time has shown, unfortunately it was the other way around: instead of saving, he condemned thousands of future children to poverty, hunger, rationed food: the food rations were introduced in 1962, when he was minister (7 months before that, at an international conference, he boasted that Cuba will become self-sufficient for food and will have 10% economic growth per year). I grew up in a communist country in the 80’s, a country that had food rations. Trying to imagine what would have been to live like that all my life helps me to understand, easier than most, the actual negative impact Che Guevara had on Cuba.

Ernesto Guevara de la Serna was born in a good, upper class, blue tribe Argentine family that claimed to have aristocratic origins. The family lost its wealth in time, but not its social status. His parents, especially his mother, were cultured people, nonconformist for their time and place, irreligious, left-leaning, supporting the left during Spanish civil war, democrats, opposing Peron’s dictatorship. They had a messy but welcoming house, full of books everywhere, on every piece of furniture.

Che’s mother was an unusual woman, the first in their city to wear pants and to drive a car. From this woman, Che inherited, apparently, both his courage and his non-conformism, but also the asthma that tortured him all his life. She encouraged him to spend his bed-ridden times reading – she taught him to read and write and later she taught him French – but also to spend his healthier intervals doing sport, trips, without care, like he was healthy. Between them there was always a special relationship, kept alive even from far away, by letters, but cut brutally when Che left secretly in his Congo expedition, and she believed, like many others, that Fidel made him disappear, as it frequently happened in communist countries between No. 1 and No. 2. She died soon after, in despair.

Che’s first contact with Marxism was as a teen, just from intellectual curiosity. He was a

voracious reader. He read everything, from adventure to poetry to philosophy. Yes, he was once one of us, did you know? Not exactly a nerd – too successful with women for that and also playing sports, in the rugby team. But he also played chess and read chess books, read philosophy and wrote his notes in his philosophy notebooks, and later will abandon a budding research career to join the revolution. He did research and published 2 papers about allergy, first as a medical student (with his professor as first author), then as a young doctor, in Mexico. He also loved (and sometimes wrote) poetry.

Besides intellectual curiosity, there was rebellion. As a little boy, he threw firecrackers through the window of some neighbors who had a dinner party. After he grew up and couldn't do that anymore, because he was himself present at such upper class dinners, with his family, he noticed that some well chosen ideas and quotes from his readings had almost the same effect. To discuss religion, for example, he used Marx, but also Nietzsche.

He was just trolling actually. His first contact with Marxism didn't convince him. In high school and later in medical school, he wasn't considered a communist, neither by himself nor by the others. He wasn't actually much interested in politics. He was interested in books, travels and women. In the last domain, like in his lectures, he showed the same voracious appetite and taste for diversity: white women, black women, Indian women, teen girls, older experienced women... He didn't yet have a political worldview, but he had an instinctive, spontaneous rejection of everything upper class, of the unwritten codes of the world he grew up in. He defied this world through his ideas and his looks (he pushed his neglect of looks close to neglect of hygiene), waiting for the moment when he could break with it completely, leaving, going far away, to discover another world, another kind of people.

Until then, he used his vacations to travel: through Argentina on bike, on the ocean, temporarily employed on a ship... In his last (prolonged) vacation, he made his famous motorbike trip. Then, he completed his medical studies and immediately left the country. First stop was Bolivia, who just had a revolution and was doing agrarian reform. At first, he liked this – it echoed his yet vague feelings of a need for social reform, feelings aroused in his former motorcycle trip by his first contact with poverty and suffering. That, until he saw a Ministry clerk spraying the Native American peasants, who came to receive land, with insecticide; disgusted by the attitude of the new revolutionary government toward common people, he leaved, going further...

Until he reached another country in the fever of revolution: Guatemala. Here he lived interesting times: he saw the contra-revolution orchestrated by the CIA, that replaced an elected president, Arbenz, with a dictator, Castillo Armas, to put an end to the agrarian

reform and nationalizations that touched American interests. He already flirted with Marxism when reaching Guatemala, but after he saw the planes shooting civilians, after he offered to fight for the revolution but was sent to a hospital, as a doctor, after he saw the demise of the president under USA pressure and the repression that followed, the arrests (including his future wife, Hilda), he came from this experience completely radicalized, full of powerless revolt against the Americans and convinced that the solution for Latin America's problems is to be found behind the Iron Curtain. He left for Mexico, where he met Fidel Castro, who invited him to join his revolution (because he needed a doctor).

The Cuba landing was a disaster. They were discovered immediately by the army and destroyed. From about 80 fighters who came with the ship, only about a dozen survived, dispersed in the mountains. Che himself was shot in the neck and felt he was dying, but survived. And still, there, in the mountains, the minuscule rebel group survived, grew more and more, fought and, after 2 years, finally won. How it was possible? With the help of the local peasants, who hid them, fed them and joined their fight, and also due to the urban people from the city movement, mostly students and young professionals, who, risking their lives, constantly send them weapons, money and food. The peasants wanted land, the urban young people wanted freedom – neither had any idea that they were fighting for the victory of communism!

Batista was a very corrupt dictator, who transformed Cuba in a brothel and had ties with the American Mafia. He was despised by almost everybody. Neither the soldiers nor the officers wanted to die for him, considering that the rebels usually freed the prisoners after taking their weapons and even treated the wounded enemy soldiers (even Che had to take care of them, as a doctor).

Although he was invited to join as a doctor for the rebels, Che preferred the role of the soldier. Just after arrival, he abandoned a first aid kit to take some ammunition instead. Soon Fidel noticed his unusual courage, so he made him "comandante", giving him the command of a column. Soon he also made himself noticed by his toughness towards his troops. He was so courageous and wanting to do great deeds that Fidel had to give him written order to abstain from suicidal actions, to not risk his life without good reason. On the contrary, Che created in his column a "Suicide Squad" for the most dangerous missions, and his rebel soldiers competed among them to join it.

After the victory, Fidel made him manager of the National Bank - although his only tie with banks, before victory, was that he wanted to rob one for the revolution, but the other rebel leaders dissuaded him - and later Minister of Industry. But Che's nature wasn't made for office work, for working with numbers. Depressed, he convinced Fidel

to send him secretly in Congo with a group of black Cuban soldiers, to help the rebels who fought there against neocolonialism.

The Congo rebels weren't exactly enthusiastic that they received a "white savior". According to Che, the situation there was a complete disaster. The rebel leaders fought among themselves and partied in the cities with the money from supporting socialist countries and didn't visit the front for months. An officer, an educated man, speaking fluent French, claimed that the bullets can't harm him due to a Congolese magic named *dawa*. The soldiers had good weapons from the socialist friendly countries but they didn't know to use them correctly and discharged them involuntarily, harming each other. During a battle, the Congolese ran away, letting only the Cubans to fight in their stead. After the battle, they punished the witch doctor because he didn't do their *dawa* well. Of course, they lost and Che and his Cubans had to retreat from Congo, vanquished.

But Che couldn't come back vanquished. The fact that he had in Cuba 5 children, the smallest one newborn when he left, wasn't enough to counterbalance his pride - sometimes the aristocrat was still visible in the communist. He needed a new battlefield to redeem himself. So, with a troop of the best veteran guerilla fighters, he went to Bolivia, hoping that, from there, he will spread the revolution in all South America, his dream being a united communist Latin America.

Why did the guerilla tactics that succeeded in Cuba lose in Bolivia? Maybe because in Bolivia they lacked the support of the peasants, who saw them as the foreign invaders they were. He only succeeded to recruit a few workers from the cities, communists disavowed by their own party leadership, but not a single peasant. The Native American peasants, who received land at their revolution, as I said, looked at these white armed bearded aliens with fear and hostility and avoided them as much as possible. On top of this, Che really had bad luck. They became totally isolated after their radio transmitter broke and Tania, their contact with the Cuban spy network, had to stay with them after a reckless mistake. Maybe she was attracted by the guerilla life or maybe she was attracted by Che, although the book doesn't claim that they were lovers. (But Che had an extraordinary charisma, he fascinated not only women but even men sometimes, once he received a declaration of love from a Soviet official!) And after the Bolivian army discovered the hiding place where they kept their documents, food and medicine, Che couldn't control his asthma anymore and the illness subdued him, transforming him in an invalid who could barely walk. What followed was a long agony...

The book is well documented and rich in details, sometimes captivating like a good

novel, other times boring due to that multitude of details that offered context but interrupted the story.

In conclusion, to paraphrase Steven Weinberg, for good people to do evil, that takes ideology. But, as someone quoted in the book said, I can't help admiring him. In spite of everything.

Consciousness Explained by Daniel Dennett

Introduction

I have always been fascinated by what philosophers call the *hard problem of consciousness*. This refers to the problem of explaining how our subjective experience can arise from a physical system such as the brain. If the brain is ultimately a collection of atoms following the laws of physics, couldn't it function perfectly well without there being anyone "at home"?

Over the years I've read many books on the subject, but none of them seemed to address this fundamental difficulty. On the one hand there were mathematically oriented books filled with many entertaining insights into the nature of recursion and abstraction but lacking content on actual human brains. On the other hand were the cognitive science and psychology books which investigated topics such as human behavior and language in great detail but, in my opinion, failed to provide a convincing account about why we ultimately feel anything.

I recently finished reading *Consciousness Explained* by Daniel Dennett and as a result, for the first time I feel as though I have made progress in understanding the hard problem. In my opinion, the defining feature of the book is that Dennett isolates a deeply held intuition many of us share which makes the "hard problem" so confusing. He then proceeds to weaken this intuition via careful examination of some of our common subjective experiences.

Contrary to the book's title, I still have no idea how consciousness really works. But I now have a lot less trouble accepting that *in principle* it could be explained in terms of simpler components such as neurons or other biological machinery. It has slowly inched from the realm of eternal mystery towards that of a wonderfully complicated scientific problem.

In this post I will describe the key idea that I learned from the book, and show how it cleared up, or at least reduced, my confusion about the hard problem.

The Cartesian Theater

What Makes Consciousness So Confusing?

Suppose that there be a machine, the structure of which produces thinking, feeling, and perceiving; imagine this machine enlarged but preserving the same proportions, so that you could enter it as if it were a mill. This being supposed, you might visit its inside; but what would you observe there? Nothing but parts which push and move each other, and never anything that could explain perception. –
Gottfried Wilhelm Leibniz

In this post I'm making the materialistic assumption that all aspects of the mind, including consciousness, can in principle be explained in terms of the physical laws governing the matter in our brains. In particular, I take the reductionist view that it is possible to create a conscious system out of simple unconscious building blocks. Indeed, the human brain is built out of electrons, protons and neutrons which are, to the best of our knowledge, completely described by quantum mechanics.

The notion of building complex systems out of simple building blocks is not new to us. It is amazing that we can build airplanes, but we do not worry about how it is possible to build a flying machine out of screws and bolts that can not fly on their own. We are all familiar with computer processors that can perform bafflingly complicated mathematical calculations which are built out of billions of simple transistors.

Perhaps more relevant is our understanding of life. We still do not understand life well enough to produce a living organism from scratch and there is still a debate as to how life should be defined. But there are no longer many vitalists arguing that there is some fundamental aspect of life that can never be explained in terms of the biochemistry of molecules such as DNA and proteins.

But when it comes to consciousness, it is hard to accept that it could ever emerge from the mindless interactions of simple components. And even when we do try to embrace that perspective, we end up getting tangled up in various paradoxes and never ending philosophical arguments. Why do the principles that we so readily apply to other unimaginably complex systems fail to satisfy us when applied to consciousness?

The reason, according to Dennett, is that our intimate familiarity with consciousness produces strong intuitions which directly contradict the reductionism described above. The paradoxes and confusion that we encounter while thinking about consciousness are the result of simultaneously entertaining mutually contradictory ideas. A major contribution of *Consciousness Explained* is to pinpoint the precise intuition that is at odds with reductionism. Once it is out in the open, we will proceed to investigate some of our common conscious experiences and demonstrate that this intuition is not as obvious as it appears.

The Cartesian Theater

Our intuitive models of the brain start with what appear to be peripheral components such as the retina, optical nerve, visual cortex, motor control and memory. We tacitly assume that these are channeled, or re-presented, through some type of central control center where thought and experience happen. In other words, we think that our brain mechanically processes inputs or execute outputs and that we only become aware of these phenomena when they pass through our “turnstile of consciousness”. Dennett calls the idea of a central control center *Cartesian Materialism* and refers to the control center itself as the *Cartesian Theater*. According to this theory, there exists a specific region deep in the brain where “it all comes together and consciousness happens”. For example, the theory claims that we are not aware of light entering our eyes until an image of our surroundings is presented to the Cartesian Theater.

This intuitive notion of consciousness is clearly at odds with reductionism and few people today would argue that it is literally true. For starters, there is the issue of infinite regress which comes up when you try to explain the Cartesian Theater itself. If the theater is simply composed of atoms then nothing is gained by assuming its existence. But if it cannot be explained in terms of simpler components then we must ultimately choose between the theater being either a single atom or some exotic form of matter than has never been detected, neither of which are particularly compelling.

Despite the Cartesian Theater’s logical flaws, it plays a major role in our intuitive understanding of the mind. As we will see later, our intuitive embrace of the theater is the foundation of many confusing paradoxes and thought experiments surrounding the study of consciousness.

The plan for the rest of this post is straightforward. We start by describing Dennett’s *Multiple Drafts* model which is his alternative to the Cartesian Theater. We will then revisit some of the more vexing paradoxes of consciousness. In each case, we will expose the Cartesian Theater as the underlying source of confusion, and show that a careful inspection of everyday

experience favors the Multiple Drafts model, even at an intuitive level. Once this is accomplished, the clash between intuition and logic disappears and the force of the paradox is greatly diminished.

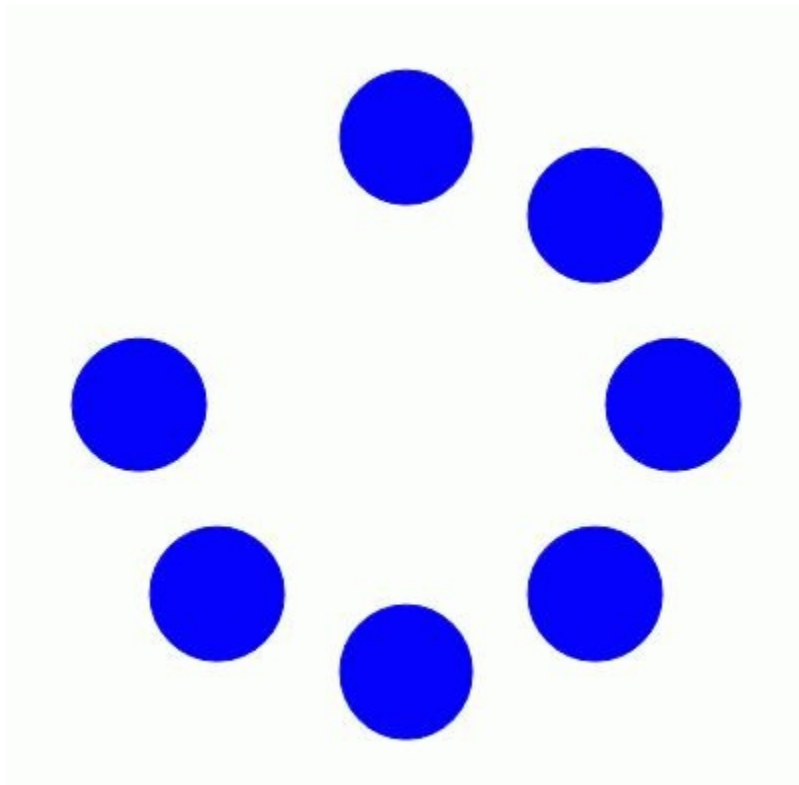
The Multiple Drafts Model

In this section we present Dennett's alternative to the Cartesian Theater which he calls the *Multiple Drafts* model. The key idea is that the brain is composed of many smaller and domain specific subsystems, all operating in parallel and each performing a simple task. The subsystems are connected and constantly process incoming signals and generate outputs, resulting in a perpetual flux of mental activity which occasionally converges on a stable pattern. The "drafts" refer to the different interpretations of our sensory input that are constantly being promoted by our nervous subsystems.

At first glance the Multiple Drafts model is at odds with our intuitive understanding of the brain. It feels like our mind proceeds in an orderly progression starting by collecting inputs which are then presented to our internal "observer" and "thinker" which makes a judgment and ultimately executes a motor command. In contrast, the Multiple Drafts model dispenses with the centralized command center and replaces it with a distributed collection of low level systems, each pursuing a simple objective until a broad consensus is ultimately reached.

In the sections that follow, the burden will be upon us to demonstrate that the Multiple Drafts model is a better description of our daily experience, even at an intuitive level. That being said, we will end this section with a simple, though perhaps contrived, example which highlights the difference between the two models.

The *phi phenomenon* (as seen on tv) is the apparent motion one observes when nearby signals alternate at high frequency. For example, consider the following image:



Source:

https://en.m.wikipedia.org/wiki/Phi_phenomenon

The blue circles in this image are just blinking in place with an offset of 150ms, but we observe a moving white dot which seems to exist even in the empty space between them.

Let us try to explain this with the Cartesian Theater model. According to that model, the change in color of each circle travels from the screen, to our eyes, through the optical nerve and ultimately is presented to the theater at which point we become conscious of it. But consider the state of the theater right after one of the dots turns white but while it's neighbor is still blue. Since there is no reason to expect the neighbor to blink, why would the audience in the theater decide at that moment that there is a white dot moving between them?

One possible explanation is that the audience prudently reserves their judgment for at least 150ms until the neighboring dot has a chance to change color. This theory can be tested by asking subjects to press a button as soon as they see a specific dot change it's color. It turns out that the response time minus the time it takes for signals to travel through the necessary nerves is significantly less than 150ms. In other words, there is not time for the theater audience to sit around and wait for the next dot to blink. One can wriggle out of this by all sorts of decidedly non intuitive theories of memory erasure and so on but our goal was never to use the

phi phenomenon to logically disprove the Cartesian Theater - we've already seen that it logically contradicts our reductionism assumption. Rather, the point is to show that our *intuitive* embrace of the theater is not as intuitive as it seems.

The key issue with the Cartesian Theater in this example is that it imposes a strict order on the events entering our conscious experience based on the time it takes signals to reach a specific point in the brain. It implies a single instance in time in which a signal transitions from an *unconscious* signal originating from the eye to a *conscious* one as it passes the threshold of the theater. This is why attempts to explain experiences such as the phi phenomenon which involve non-linear (in time) reactions to sensory input are bound to get us in trouble.

On the other hand, the Multiple Drafts model easily handles the phi phenomenon without relying on explanatory contortions. According to this model, the change in the first dot generates multiple "drafts" of mental activity corresponding to a stationary dot, a moving dot and many other possibilities. The lack of supporting evidence for the non stationary drafts allows the stationary one to stabilize and ultimately trigger a button press. When the next dot blinks the suppressed moving draft takes over and this new dominant pattern of mental activity ultimately causes the person to say (out loud or to themselves) that they saw a moving dot.

In the remaining chapters we will build our intuition for the Multiple Drafts model by inspecting some common everyday experiences. In the process of doing so we will shed light on some of the paradoxes and mysteries which typically occur in the study of consciousness.

Do We Mean What We Say?

In this section we address the following problem: Suppose you make a machine that can pass the Turing test. It can take inputs in the form of character strings and produce responses that are indistinguishable from those of a person. But the machine is clearly just mindlessly shuffling bytes without there being "anyone home". Surely this is not the same thing as a conscious person talking since we *mean what we say*. How could a reductionist theory of consciousness *ever* explain the fact that we, unlike computers, mean what we say?

Let's try and lay out the assumptions behind this argument. Our intuitive model of speech production posits a clean sequence of steps. First we decide our intent, or in other words, we decide what we want to say at an abstract level. This intent gets passed to a language specific system which generates words. These get passed to a grammatical module which fits the words

into a sentence. Finally, this sentence gets passed to the speech production system which executes a pattern of breathing and vocal cord movement.

It's possible that your intuitive model is somewhat different, but the key idea is that there is a special system that initially determines what we want to say, and that this kicks off the rest of the speech synthesis process. This high level coordinator is another word for Dennett's Cartesian Theater. The reason the problem of intent and meaning is so confusing is that it is based on the intuitive Cartesian Theater which is indeed logically incompatible with reductionism.

Before debating this any further it makes sense to take a step back and ask "Is this intuitive model correct?". Dennett claims that the answer is no. And indeed you can verify this by asking yourself a simple question: When you engage in normal conversation, do you usually know what you are going to say?

I admit that I initially found this suggestion preposterous - *of course* I do! But after observing myself during social interactions I realized that indeed words just sort of appear on my tongue and seem to be influenced mainly by previous words and elements of my environment such as other people's words. Most people I've asked about this report the same thing. The exceptions were a few people, perhaps more filtered than myself, who told me that they sound out words in their head before saying them. But this just kicks back the question up a level - they agreed that they did not know what words would appear in their heads before sounding them out. Indeed, what would be the point of talking to yourself if you already knew what you were going to say?

The takeaway is that when we really think about it, our experience does not obviously point to a top down chain of command where our abstract thoughts are systematically converted into utterances. Rather, certain aspects of our experience are consistent with a chaotic marketplace of neural subsystems constantly forming and reforming coalitions.

We can find further evidence for this perspective by considering exceptions to the typical process of speech generation.

Sometimes a person will say something for no particular reason other than the fact that they like the way it sounds. For example, ever since I started watching Brooklyn 99 I find myself reacting to many situations with Jake Peralta's patented catchphrase "Noice", even when the underlying meaning of "nice" is not strictly warranted. On a more cynical note, a few weeks ago I asked a receptionist how much longer it would take to be seated and when they answered "15 minutes" I

immediately said to myself “can save you 15% or more on car insurance”. These phenomena fit the Multiple Drafts hypothesis quite well. There are multiple subsystems in our brain trying to trigger a system wide response and sometimes the sillier “drafts” end up winning control.

A more extreme and unfortunate example is a condition called *jargon aphasia*. People with this condition frequently make nonsensical word choices, sometimes uttering completely incomprehensible sentences, despite performing well on other measures of intelligence. It appears as though they produce speech without *any* high level directive.

Many of the principles that apply to speech also hold for our internal monologue. It too does not follow a strict top down policy and can frequently be influenced by common templates and word patterns that we pick up from others. This is at least the way it feels to me when, for example, I talk myself through the steps of a well known procedure.

In summary, our difficulty with believing that a machine could reproduce all aspects of human speech is based on an intuition that speech is initiated from a single source of meaning and intent - the Cartesian Theater. This intuition is not compatible with a piece of machinery that can be reduced to simple components, each devoid of meaning. Our approach to resolving this conflict is to challenge the Cartesian Theater by providing intuitive support for the Multiple Drafts model.

Awareness

A key aspect of our conscious experience is that we do not just mechanically react to sensory inputs. There is clearly a distinction between *reflexive* actions such as recoiling from a hot object and *deliberate* activities like drawing a picture or solving a math problem. The latter seem to involve an inner *observer* or *witness* to the proceedings. How could a computer *ever* perform a deliberate action? Isn't anything a machine does reflexive by definition?

It may not surprise you to hear that this question is also rooted in a profound intuition for the Cartesian Theater. Indeed, our intuition is that reflexes exploit hard coded circuits connecting sensory inputs to motor control but deliberate actions require these inputs to pass through the “turnstile of consciousness” so that the audience in the Cartesian Theater may devise an appropriate response.

In this section we will analyze these two fundamental types of actions and see that the difference is not as clear as it may seem. As usual, the goal is not to provide a logical proof of anything, but rather to weaken our intuition for the Cartesian Theater just enough so that we can begin to entertain alternatives such as the Multiple Drafts model.

The Unconscious / Conscious Boundary

One strong source of intuition for the Cartesian Theater lies in the apparent distinction between conscious and unconscious actions. An example of this distinction is the phenomenon of an object “hiding in plain sight”. It is possible for us to look directly at an object such that the light reflecting from it passes into our eyes and down our optical nerve but still not be *conscious* of it. At some point something happens and the object passes across some sort of boundary into the realm of consciousness.

We can study this boundary by noting that awareness seems to be necessary for following instructions. Consider the following command: “Raise your right hand if you unconsciously observe a blue square”. It would appear as though this command makes no sense. How could you decide to raise your hand without being consciously aware of the square? However, it is actually possible to fulfill this strange request. All you have to do is practice by sitting in front a computer flashing various shapes and raising your hand when you see a blue square. Over time you will get better and better at this until you can raise your hand in response to the square before you even know what’s happening.

This example seems contrived, but far from being the exception it seems to be the rule. Learning a new skill involves a gradual transition from the realm of conscious deliberations to reflexive and automatic reactions. This phenomenon is evident in physical activities such as playing a piano or shooting a basketball but it is also a key feature of learning mental skills. For example, when we learn multiplication tables in school we eventually get to stage where the phrase “3 times 7” causes us to reflexively respond with “21” - either out loud or to ourselves.

Here is another example which may be of practical interest to undergraduate students. Exams in large entry level mathematics classes are typically graded by assigning each question to a different graduate student. Each grad student will grade hundreds of responses to their question over the course of a few hours. In my experience, which has been confirmed by some of my friends, the grading process starts slowly as care must be taken to find the spot on the page containing the answer and verify the student’s reasoning. However, after looking through around

100 exams I sort of “get in the zone” and can predict a student’s grade after staring at their worksheet for about two seconds.

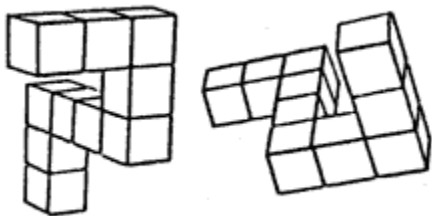
We have seen some examples where a conscious activity can slowly turn into an unconscious one. The converse is possible as well. For example, we can train ourselves to become aware of certain sounds and tastes that we currently ignore. A big source of examples is learning a foreign language. If you immerse yourself in a foreign country the language will initially sound like an indistinguishable stream of sounds. Simply learning a single word of the language will cause you to start hearing it everywhere. The same goes for learning to recognize subtle pronunciation differences.

As another example, we can learn to recognize flavors in foods like wine, beer, chocolate and coffee. With practice we can become aware of flavors that we otherwise would not have know existed.

In summary, many aspects of our learning experience are not easy to reconcile with the strict turnstile of consciousness imposed by the Cartesian Theater. How about the Multiple Drafts model? In this model, we become aware of sensory input if that input triggers a self sustaining pattern of mental activity. The strength and duration of the pattern determines our awareness level. When we learn a new skill, we create connections and frameworks in the brain that make it easier for certain patterns to persist.

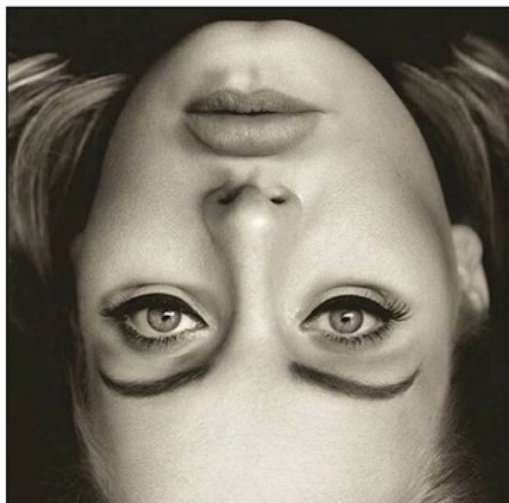
The Mind’s Eye

One strong source of intuition for an internal observer is the phenomenon of the *mind’s eye*. To be concrete, consider the following *mental rotation* question [Shepard (1971)]¹. Are two figures below the same except for their orientation?



As you attempt to solve these sorts of puzzles, it may feel like you are rotating the shapes in your mind's eye. This intuition is supported by the fact that the time it takes to answer these type of questions has been shown to be linear in the degrees of rotation that are required. This seems to be good news for the Cartesian Theater - we solve mental rotation problems by projecting a rotating version of the reference shape in the theater and waiting until it looks the same as the comparison shape.

Things start to get interesting when we consider simple variations of this experiment. Perhaps the most famous is the *Thatcher Effect* where our random assortment of blocks is replaced with a human face. The result is that local changes to the face are significantly harder to detect when the face is turned upside-down. Here is an example with Adele:



[Source](#)

In this example, Adele's eyes and mouth have locally been flipped upside-down. This is obvious, and quite disturbing, when we look at the image on the right, but less so for the rotated version on the left. If we truly possessed the ability to rotate images in our mind's eye, shouldn't we be able to mentally rotate the image on the left by 180 degrees and experience the same reaction that we do to the one on the right?

Another interesting line of research studies the relationship between the object's familiarity and the mental rotation speed. In general, we rotate familiar objects significantly faster than arbitrary ones. E.g, we can rotate letters of the alphabet more quickly than random squiggles. A cool demonstration of this was given by Sayeki (1981) ² who showed that even providing the subject

with a familiar *interpretation* of an abstract shape reduced the time they needed to solve a mental rotation problem by as much as 2 seconds! For example, in the case of the block image above they told subjects to imagine that the left image represented a person sitting on a chair and sticking out their left arm.

One takeaway from these results is that solving mental rotation problems is not a simple matter of turning a literal picture around in our head. Rather, it appears as though we are also interacting with abstract mental structures (such as a face or chair) that we associate with the reference image. As usual, this is not irreconcilable with the Cartesian Theater but clearly our initial intuitions do not tell the full story.

Qualia

The philosophical notion of *qualia* frequently comes up in the context of understanding consciousness. The idea is that objects in the world seem to possess certain subjective *qualities* which can not be explained in terms of their physical attributes. A common example is the quality of the color red. Consider this beautiful red square:



[Source](#)

Proponents of qualia would note that even if you perfectly describe the wavelength of the light coming off the screen, the electrochemical reactions it triggers in your retina and the neural activity induced in your brain, there is still some essential *redness* which has not been explained. Since this redness is clearly part of our consciousness, doesn't this mean that physical explanations of consciousness are doomed to failure?

Due to the fact that the very definition of qualia is debated by philosophers, in this section we will focus on a famous thought experiment proposed by Frank Jackson³ which goes as follows:

Mary the color scientist knows all the physical facts about color, including every *physical* fact about the experience of color in other people, from the behavior a particular color is likely to elicit to the specific sequence of neurological firings that register that a color has been seen. However, she has been confined from birth to a room that is black and white, and is only allowed to observe the outside world through a black and white monitor. When she is allowed to leave the room, it must be admitted that she learns something about the color red the first time she sees it — specifically, she learns what it is like to see that color.

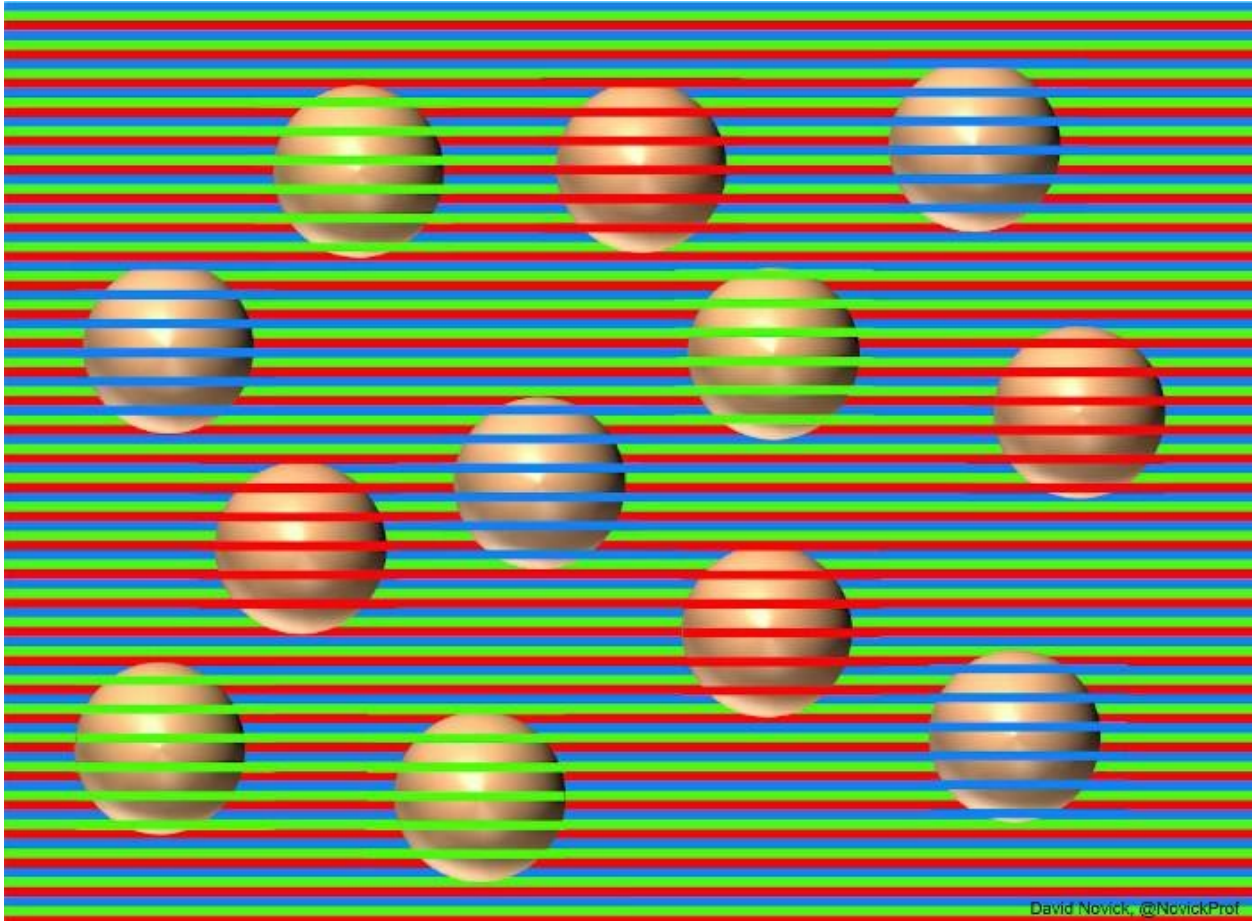
The takeaway is that there is a subjective aspect to color that can not be captured by physical facts. The only way to know *what it feels like* to see color is to actually see it. We can all imagine Mary seeing a red rose for the first time and exclaiming “So *that’s* what red looks like!”.

In this case we will confront the issue directly rather than following our usual program of tracing it back to a lingering belief in the Cartesian Theater.

According to Dennett, the fundamental problem with this thought experiment is that we take it for granted that even after knowing *every physical fact* about color, Mary will still be surprised when she steps outside of the room and sees red for the first time. In Dennett’s words, we are “mistaking a failure of imagination for an insight into necessity.”

He claims that we are greatly under-estimating what it would take to really know *everything* about the color red. As a first approximation, maybe this involves knowing that the wavelength of red light is about 680nm and that this light triggers a specific set of cells in our eyes. Everyone agrees that if all Mary knew was a perhaps embellished version of this description then upon actually seeing red she would be quite surprised indeed.

But what really *is* color? It is now well known that our color perception of an object is influenced by many factors including, but not restricted to, the wavelength of light it reflects. As a cool illustration of this, consider the following image:



[Source](#)

It may surprise you to hear that all of the balls in the image have exactly the same RGB values. You can verify this by covering up the lines (the source contains an animation of the lines being added and removed) or zooming in until the lines become far apart. From this example we learn that the color we associate with an object is influenced by its environment.

Another demonstration of the precarious relationship between color and wavelength is the absence of a wavelength corresponding to the color brown. Look at a rainbow and note that there is no brown to be seen! There is clearly *some* visual property shared by chocolate, chipmunks and violins but wavelength alone can not describe it. I'd recommend checking out [this video](#) if you want to learn more about brown.

Finally, phenomena such as [The Dress](#) show that even people with perfectly healthy retinas may not be able to agree on what colors appear in a given picture. In fact, even an individual may experience color differently depending on the season of the year! ⁴

From these sorts of examples we conclude that it is not possible to define color in the abstract without referring to the specific human brain that is observing the world. A *complete* description of the color red only makes sense in the context of a single person at a particular point in time, and involves specifying a precise mental pattern that is triggered by various objects and images.

What does all of this mean for Mary? Well, in the process of learning everything there is to know about the color red she must have discovered the exact pattern of mental activity that firetrucks and roses would trigger in her brain. She could use this knowledge to create a machine that would stimulate her brain in that precise fashion. Having done so, it is reasonable to expect that she would not experience anything new when she steps out of the room.

You may object that creating a machine brain interface is not allowed. But why should be Mary's brain modification technology be restricted to old fashioned tools such as words and diagrams? It is possible that language is simply not rich enough to convey something as subtle as redness. Even with this restriction, it is also possible that a person with Mary's capabilities could devise a clever string of words that would unlock the desired neurological response.

Philosophical Zombies

A philosophical zombie is a being which behaves exactly like a human but has no subjective experience. It is easy to imagine such an entity. For example, maybe it is a robot controlled by a massive array of GPUs running the latest and greatest machine learning algorithms. On the outside it would engage in typical human activities such as watching Netflix and tweeting stuff. But on the inside it would be cold and dark - just electric impulses shuffling through wires without anyone "at home".

The existence of such a zombie seems plausible, but doesn't this imply that no physical description of the brain could ever bridge the gap between a mindless zombie and a conscious human being?

Dennett's answer to this question is simple: No, philosophical zombies are *not* possible.

One of the reasons we think we can imagine them is our stubborn intuition for the Cartesian Theater. Following this intuition, one makes a zombie by replicating all of the machinery in the brain *except* for that special area containing the theater.

But as we have seen, such a theater is logically inconsistent with reductionism. Furthermore, even the theater's intuitive appeal is diminished after careful inspection of some common subjective experiences.

Another reason for the seductive nature of the zombie argument is that we greatly oversimplify what it would take to build one. For example, it must be possible for you to collaborate with this zombie on a complex project for many years without you noticing that anything is off. Such a machine would be significantly more advanced than anything we could build today and it is premature to assume that we know anything about it at all. We certainly are not in a position to speculate about the nature of its consciousness.

Conclusion

I hope I have been successful at explaining why Dennett has changed the way I think about consciousness. The key realization was that the philosophical knots I used to tie myself in while thinking about the subject could be traced to my simultaneous belief in reductionism and intuition for the Cartesian Theater.

The next step was to consider a possible alternative to the theater such as the Multiple Drafts model.

Then came the real work of shifting my *intuition* from the pernicious Cartesian Theater to this new model. This was done by carefully inspecting key aspects of consciousness such as speech, thought and awareness which defy simple theater theoretic interpretations.

To conclude, here is a quote from the book in which Dennett provides a "thumbnail sketch" of his theory:

There is no single, definitive "stream of consciousness," because there is no central Headquarters, no Cartesian Theater where "it all comes together" for the perusal of a Central Meaner. Instead of such a single stream (however wide), there are multiple channels in which specialist circuits try, in parallel pandemoniums, to do their various things, creating Multiple Drafts as they go. Most of these fragmentary drafts of "narrative" play short-lived roles in the modulation of current activity but some get promoted to further functional roles, in swift succession, by the activity of a virtual machine in the brain. The seriality of this machine (its "von

Neumannesque” character) is not a “hard-wired” design feature, but rather the upshot of a succession of coalitions of these specialists.

The basic specialists are part of our animal heritage. They were not developed to perform peculiarly human actions, such as reading and writing, but ducking, predator-avoiding, face-recognizing, grasping, throwing, berry-picking, and other essential tasks. They are often opportunistically enlisted in new roles, for which their native talents more or less suit them. The result is not bedlam only because the trends that are imposed on all this activity are themselves the product of design. Some of this design is innate, and is shared with other animals. But it is augmented, and sometimes even overwhelmed in importance, by microhabits of thought that are developed in the individual, partly idiosyncratic results of self-exploration and partly the predesigned gifts of culture. Thousands of memes, mostly borne by language, but also by wordless “images” and other data structures, take up residence in an individual brain, shaping its tendencies and thereby turning it into a mind.

1. Shepard, Roger N., and Jacqueline Metzler. “Mental rotation of three-dimensional objects.” *Science* 171.3972 (1971): 701-703. ↵
2. Sayeki Y (1981) “Body analogy” and the cognition of rotated figures. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition* ↵
3. Jackson, Frank (1982). “Epiphenomenal Qualia”. *The Philosophical Quarterly*. 32 (127): 127–136 ↵
4. Welbourne, Lauren E., Antony B. Morland, and Alex R. Wade. “Human colour perception changes between seasons.” *Current Biology* 25.15 (2015): R646-R647. ↵

Creative Evolution by Henri Bergson (1907)

A radical view of evolution, founded on a very different understanding of time and knowledge.

The full text of this book can be found at: <http://www.gutenberg.org/files/26163/26163-h/26163-h.htm>

Ch. I

Bergson starts with a description of what it is like to experience time.

"My mental state, as it advances on the road of time, is continually swelling with the duration it accumulates." [p.8]

The present contains all of the past. Even if it's not explicitly remembered, it builds on our character. The future is larger than the present and so it cannot, in principle, be predicted - a predicted future is smaller than the present. Anything which builds on itself has this sense of time - psychology, species, life as a whole, the entire universe.

Mechanistic science has time as a parameter. It selects some simple system and artificially isolates the system from the rest of the world. Because it is simple, its future can be exactly predicted by its present - using a differential equation. This is an insufficient understanding of time to describe anything that really builds on itself / has duration / endures.

"Science can work only on what is supposed to repeat itself -- that is to say, on what is withdrawn, by hypothesis, from the action of real time. Anything that is irreducible and irreversible in the successive moments of history eludes science." [p.36]

Both the mechanistic and finalistic views come from humans' need to design tools. In building a tool, we think of it as a collection of simple interacting parts - the mechanistic view of science. The tool also has a purpose (an end cause) - the finalistic view of teleology. Neither is an accurate depiction of time. They can both be explained by tendencies of the human intellect which help us survive.

For life, we can't separate out a simple system because things have identity. An animal exists in a sense that a rock does not because the animal's parts are bound together and it can only exist as a whole. Along with this tendency towards identity, life has a contradictory tendency towards reproduction. He describes all reproduction as a form of budding - splitting off a little piece of a creature, which can then exist on its own. This binds all life together as a current of continuous progress. The entirety of life has duration - not just individuals. Organic evolution resembles the

evolution of consciousness.

Next is a description of various contemporary (1907) models of evolution:

- Darwinism - mechanistic - accumulations of random variations - negative selective pressure - gradual, infinitesimally small changes.
- Punctuated equilibrium - similar to Darwinism - large, abrupt changes.
- Lamarckism - certain organs are strengthened or weakened by use or disuse - these changes are inherited.
- External environment - imposes its form on the creatures.
- Teleology - biology has an end goal - all life is one.
- Bergson's own view - original impetus of all life - elan vital (more on this later).

I had previously had a much simpler view of the contending theories.

Bergson criticizes all of these.

Teleology has an inaccurate notion of time. The end can't be known from the beginning because the future is bigger than the past. The unity of life is in the past, not the future.

Bergson's criticism of the external environment is easier to explain using post-Shannon language: environment transmits less information than the organisms gain.

Bergson then reviews the available evidence for Lamarck vs Darwin (gradualism) vs punctuated equilibrium. The main evidence for Lamarck is alcoholism and an experiment where you cut a nerve of a guinea pig & its descendants have epilepsy. Giraffes are not mentioned at all. Despite this, most, if not all, acquired traits are not inherited. He doesn't have a strong opinion on gradualism vs punctuated equilibrium. Bergson's criticism of all three (Lamarck, Darwin, punctuated equilibrium) is that they don't explain similar structures arising in different evolutionary branches. Convergent evolution can explain some of these. The main example he uses is that vertebrates and cephalopods have the same eye structure.

Ch. II

Life has an original impetus / tendency, which drives it to increasing diversity. Along the way, life develops many (subsidiary?) tendencies. Although the ideal might be to achieve all of them at once, that is typically unrealizable. Different branches of life increasingly focus on one at the expense of others. The tendencies are complementary and antagonistic. Groups should be defined by their tendencies in a certain direction, not the presence of measurable properties.

Adaptation to the environment is a necessary constraint, but it does not explain the driving force.

The first main division in life is between the impetuses of the two kingdoms: animals and plants.

Plants have tendency to gather and store large amounts of potential energy; animals have a tendency to release potential energy explosively/abruptly. Storing potential energy is inherently tied to creating organic matter out of mineral elements and to immobility (both can be referred to using 'fix'). Releasing potential energy is inherently tied to movement and consciousness, the means of directing movement.

Within plants, the next big division is between fixing nitrogen and fixing carbon. Bergson correctly recognizes that fungi don't fit well in this description, but explains that as a regression or torpidity of plants, not as a kingdom with its own impetus.

Within animals, the main division is between an impetus towards instinct and an impetus towards intelligence. Many animals have gotten stuck in torpidity by seeking defense in a hard shell, instead of further pursuing movement - echinoderms, most mollusks, some arthropods, early fish. Intuition has reached its fullest development in ants. Intelligence has reached its fullest development in man. Although both are always intertwined, intuition is stronger most of the time, even among vertebrates.

Intuition is the capacity to use a ready-made tool (a body part) for a particular job. Intelligence is the capacity to design tools for any job. Humans should be considered *Homo faber*, defined by our ability to make tools.

Consciousness is the difference between thoughts and actions. The unconsciousness of an ant is very different from the unconsciousness of a plant or stone. The latter has no thoughts. The thoughts of the former are exactly realized in action. In colloquial language, a 'conscious choice' is one in which we are thinking about other alternatives as well as the choice we are making.

"It lights up the zone of potentialities that surrounds the act." [p.182]

Innate knowledge for instinct is about things/matter. Innate knowledge for intelligence is about relations/form. Intuition is intimate and full, but is limited to a few objects, especially natural instruments (your own body parts). Intelligence is external and empty, which allows it to describe infinitely many objects, and construct artificial instruments. Only intelligence can ask about the essence of things, but only intuition could understand it.

Our intelligence primarily understands immobile, discontinuous, unorganized solids that can be disassembled and reassembled. It is particularly poorly suited to understand life.

Instinct is molded on the very form of life. It is impossible to distinguish where anatomy and physiology separate. Instinctual knowledge is built using the whole of our past, including the ancient unity of life. There is no reason to think instinct (the direction of one branch of life) can be entirely understood in terms of intelligence (the direction of a different branch).

Wasps which paralyze caterpillars know precisely where & how many nerve centers to strike. This knowledge is built by the species' instinctual effort. It is complicated and can't be built

piecemeal entirely by small random mutations. It does not represent habit formed by intelligence - wasps never were intelligent. Bergson considers this a form of sympathy: shared knowledge as a result of deep history. Sympathy here is not an ethical term, as this example clearly indicates.

Bergson then introduces / redefines 'intuition'. Intuition is conscious instinct reflecting on and enlarging itself. It is not just whatever your gut tells you. Art shows that intuition is possible. Bergson calls for an organized effort towards intuition, which relates to instinct and art the way science relates to intelligence and observation. Both intuition and science should continually inform and strengthen each other.

The evolution of life is a broad current of consciousness with an enormous multiplicity of interwoven potentialities. Each branch focuses more on its own particular goals rather than the work of the whole. Instinct at first seems to be the more effective strategy, since it remains focused on itself, but it is limited. Intelligence first concentrates on external matter, then gains the ability adapt to many objects, and so can awaken the potentialities of intuition. Gaining mastery by invention is more useful than the material invention itself. Humans are unique not just because we are more intelligent than animals, but because our intelligence has set consciousness free.

Ch. III & IV

There are two more chapters that I will not summarize, for several reasons:

Bergson rarely leads with his thesis statement. Instead, he tries to gradually build understanding through examples, metaphors, and contrasts. When writing this summary, I've used simpler explanations from later in the book when Bergson first begins to hint at an idea. So I've included much of the material from later in the book in the summary of the earlier chapters. The best thesis statement of the book is the last section of Ch. III.

The remaining material is more philosophical than scientific (What is the difference between order and disorder? How is modern science different from ancient natural philosophy?) and is of less interest to me personally.

I like the way this summary ends.

Other Quotes

"Perhaps even it is necessary that a theory should restrict itself exclusively to a particular point of view, in order to remain scientific, i.e. to give a precise direction to researchers in detail. But the reality of which each of these theories takes a partial view must transcend them all. And this

reality is the special object of philosophy, which is not constrained to scientific precision because it contemplates no practical applications." [p.89-90]

"[Societies of bees and ants] are admirably ordered and united, but stereotyped; [human societies] are open to every sort of progress, but divided, and incessantly at strife with themselves. The ideal would be a society always in progress and always in equilibrium, but this ideal is perhaps unrealizable: the two characteristics which would fain complete each other, ... can no longer abide together when they grow stronger." [p.105]

"Nature is more and better than a plan in the course of realization. A plan is a term assigned to a labor: it closes the future whose form it indicates. Before the evolution of life, on the contrary, the portals of the future remain wide open." [p.109]

"Life in general is mobility itself; particular manifestations of life accept this mobility reluctantly, and constantly lag behind. It is always going ahead; they want to mark time. Evolution in general would fain go on in a straight line; each special evolution is a kind of circle. Like eddies of dust raised by the wind as it passes, the living turn upon themselves, borne up by the great blast of life. They are therefore relatively stable, and counterfeit immobility so well that we treat each of them as a thing rather than as a progress, forgetting that the very permanence of their form is only the outline of a movement. At times, however, in a fleeting vision, the invisible breath that bears them is materialized before our eyes." [p.131-132]

This is my favorite analogy from Bergson. If you combine this analogy with the taxonomic hierarchy, you get a turbulent cascade model for evolution. This is incredible intuition because this understanding of turbulence hadn't been developed yet. Richardson's poem, "Big whorls have little whorls \ Which feed on their velocity, \ And little whorls have lesser whorls \ And so on to viscosity" is from 1922 and Kolmogorov's theory is from 1941.

"phantoms of ideas to which there cling phantoms of problems" [p.181]

"it is of the essence of the psychical to enfold a confused plurality of interpenetrating terms" [p.257]

"ancient science thinks it knows its object sufficiently when it has noted of it some privileged moments, whereas modern science considers the object at any moment whatsoever" [p.326]

Modern science is the daughter of astronomy; it has come down from heaven to earth along the inclined plane of Galileo." [p.331]

"For simple common sense tells us that when we are possessed of an effective instrument of research, and are ignorant of the limits of its applicability, we should act as if its applicability were unlimited; there will always be time to abate it." [p.343]

My Comments

Are there any modern (post-1970) Bergsonian biologists?

My (limited) understanding of the history of evolutionary theory has continual conflict between gradualism, punctuated equilibrium, and Lamarckism. The acceptance of the viewpoints is still shifting, based on empirical evidence - mostly from the fossil record and from genetics & epigenetics. Teleology was once widely accepted, but has gradually lost favor over many decades. Bergson ... has been completely ignored? Even if his arguments are not compelling enough to convince the entire scientific community, they are not completely ridiculous, and so should have convinced a few people to develop them further using modern evidence. My impression is that there are more biologists who are Spinozists than Bergsonians.

The particular examples are old and may have been solved or may have been forgotten. We figured out how fetal alcohol syndrome works without relying on Lamarck. Why do vertebrates and cephalopods have similar eyes? What was going on with the epileptic guinea pigs?

Bergson's description of what it is like to experience time is especially lucid. He takes the experience of time as fundamental and explains how our understanding of mechanistic time arises from that. Most scientists today take the mechanistic description of time as fundamental. Can we derive the experience of time from mechanistic time?

I've raised a similar question when talking to other physicists about the arrow of time. They typically refuse to consider it as a valid question. I couldn't elaborate on "Why do all observers experience time in one direction?" Bergson definitely can.

Bergson is sometimes described as being anti-intellectual. I think it's more accurate to say that he is anti-exclusively-intellectual. I can see how misunderstanding what he means by e.g. 'intuition' could allow you to conclude that whatever you feel is right is better than thinking through things. But what he's arguing for is a sincere / strenuous effort to improve human knowledge using both the intellect and anything we can learn through instinct. His criticism is not that we use intelligence too much but that we disregard anything coming from instinct.

Bergson develops, in detail, ideas that do not correspond to any of the words in current language. He refers to them using the closest available words ('intuition', 'sympathy', 'duration', etc.), often in a highly unconventional way. Although this makes it easier for him to be misunderstood, the other option would be to invent new words, which would make more people not understand him at all.

The ideas presented here are complex and contrary to the worldview of most scientists (and people), but Bergson does find a way to explain them. Since he is challenging our understanding of knowledge, his ideas can't always be given a precise definition. Instead, they are explained through examples, contrasts with other ideas, and vivid metaphors.

Defenders of the Truth: The Sociobiology Debate by Ullica Segerstråle

“The characters in my story are all defenders of the truth — it is just that they have different conceptions of where the truth lies.”

— *Ullica Segerstråle*

It's always risky to return to something that made a big impression on you when you were young. If at 35 that favorite movie you had at 12 isn't as good as you remember it, and the book that blew your mind at 16 is actually pretty simplistic and juvenile upon revisit, do you really want to know?

I had such worries about *Defenders of the Truth: The Sociobiology Debate*, which I first read in 2008 at the age of 24 and recently decided to pick up again. I'd recommended it on a podcast and I wanted to know whether it was as good as I remembered or if I had made a fool of myself.

I once mentioned *Defenders* when introducing the word "erisology", an umbrella term for "the study of disagreement". And that's what this book is: a long, very thorough history and analysis of an extremely complicated disagreement. Or perhaps more correctly an extremely complicated structure of disagreements involving dozens of actors and spanning several decades. In retrospect I think my own preferred approach to making sense of disagreements owes a lot to Segerstråle's work in this book. It's certainly been a goal of mine to emulate it.

"The Sociobiology Debate" gets its name from Edward O. Wilson's book "Sociobiology", published in 1975. It was a massive tome discussing the biology of behavior across many species, among them Wilson's own specialty — the social insects. That didn't provoke much debate by itself, but the opening and closing chapters did. They took the thinking from the rest of the book and speculatively applied it to humans, arguing that a future, more developed science of sociobiology would be able to explain our behavior as well.

This caused a big stir that grew to involve scholars across various subdisciplines of biology as well as academic psychologists and philosophers, plus a good helping of political activists (including a group specifically formed to fight sociobiology). The fighting went on for years in books, articles, reviews and conference talks. It morphed and mutated and brought in new

people over time, but it never really stopped. *Defenders* was published 25 years after *Sociobiology* and is itself 20 years old now. The battles have continued, more decentralised, complex and confused than before.

The author Ullica Segerstråle, a chemist turned sociologist originally from Finland and now at the University of Illinois, is likely better equipped to write this history than anybody else. Not only did she spend years doing research and interviewed all the major players about their role in the controversy, she was present at important events including meetings of the activist Sociobiology Study Group and the infamous seminar when a group of protesters poured a pitcher of water over Wilson.

Her book is divided into three parts, titled "What happened in the sociobiology debate?", "Making sense of the sociobiology debate" and "The cultural meaning of the battle for science". The first not only tells us what happened but also goes into the historical background. There's a lengthy history of the field leading up to Wilson, dipping more than once into the philosophy of science and of biology specifically. This part can be heavy to get through if you didn't come for a primer on population genetics, but it sets up the pieces for what comes later quite nicely.

After going through who said what and when for about 150 pages, Segerstråle in the second part turns to explaining the motives of the different sides and players — what they wanted, what they thought was important, what they took for granted, and the conclusions they drew from that. In the final third she returns to her own present, looks at the debate after two and a half decades to discuss its relationship to the Science Wars in the '90s, to philosophical debates about the nature of the will, and the meaning of the "Enlightenment Quest".

Wilson and Lewontin at the center

What stands out about the sociobiology debate is how many "fronts" it had and on how many levels you can describe it. From high up above there's the familiar macro-conflict between "nature" and "nurture", which here like always turns into confused, muddy complexity as you get closer. Beneath there's a set of interlocking disagreements about the proper relationship between science and politics, the duties of scientists and the criteria for what counts as good science, all energized by personal scientific ambitions, deep moral convictions, clashing personalities and academic turf wars.

Segerstråle handles this by including a large cast of characters, each with their own concerns, but zooming in and focusing on one specific conflict for a good part of the story: that between Wilson and his Harvard colleague Richard Lewontin.

Lewontin, himself a population geneticist, became Wilson's primary critic (even though he didn't necessarily want the role). This is explained by him having, at the same time, scientific and political objections to Wilson's work.

Due to their differing temperaments and academic backgrounds they didn't agree on what constituted good science. Wilson was a naturalist, who spent most of his time in the field studying animals, and considered speculation based on observation to be scientifically respectable. He believed that science progressed through bold, creative leaps into new territory where rough drafts could establish a presence to be further developed later.

His claims about humanity in Sociobiology (and in his later books "Genes, Mind and Culture" and "On Human Nature") were thus not facts demonstrated beyond reasonable doubt and not intended to be. They were plausibility arguments that would in Wilson's mind later be worked out and put on more solid ground as the science of sociobiology progressed. Therefore he saw fit to use, for instance, formulas that were known to be flawed, because they were the sort of thing a stronger version of his theories would include.

To Lewontin this was all unacceptable. His philosophy of science was that of a hard-nosed experimentalist used to dealing with molecules. He disliked speculation and is described by a colleague as extremely smart but not very creative or imaginative. As he saw it, science progressed by the gradual establishment of reliable physical facts, and so-called "science" that didn't explain things in terms of the real, concrete mechanisms involved wasn't real science. A case in point was research on intelligence and its heritability. He believed it was possible in principle to study the influence of genetics on intelligence, but only in terms of the actual biochemical effects in the brain of known, identified genes. Research that dealt with abstractions like IQ scores derived from tests and then connected this mathematical construct to hypothetical genes was definitely not serious.

Moreover, he had a particularly cynical view of what motivated researchers in the fields he considered pseudoscientific. Because properly scientific reasoning established its conclusion as an incontrovertible fact, supposedly scientific arguments that fell short of this standard — such as the speculation and plausibility appeals that made up Wilson's draft of human sociobiology — was to Lewontin scientifically worthless not just in his own personal opinion, but clearly so even to the people who engaged in it. Thus they must have ulterior motives. These motives could be careerism, or lies in the service of a regressive political agenda.

Lewontin was primed to make negative assumptions about Wilson's ideas by having in mind examples from scientists in history that he considered obvious examples of outright lying for

racist reasons. Segerstråle brings up his harsh judgment of the 19th century zoologist Louis Agassiz who he confidently declares a liar. In his eyes it was simply impossible to honestly believe something false — because it's impossible to have valid, incontrovertible, evidence for it.

At the centers of different worlds

This all illustrates how we often underestimate exactly how much and in how many ways other people's assumptions, concerns and motivations differ from our own. We use this miscalibrated understanding to draw further conclusions and build complex but often erroneous mental representations of other people's beliefs. That we also in turn overestimate *other* people's understanding of *our* motivations, concerns and assumptions and therefore misinterpret the reasons for their behavior towards us makes it even worse.

It also shows how often disagreements can't be reduced to a single issue. The conflict between Wilson and Lewontin wasn't purely scientific nor purely political. Scientists in different fields may disapprove of each other's methods but typically prefer to keep quiet to making a stink about it. But if politics come into play the equation changes. Lewontin was a self-described Marxist and radical who believed that bad science could easily be used by actors trying to justify the social status quo, in a way that good, *correct* science could not. In this way his scientific and political beliefs were intimately connected.

Given some core disagreements, and the fact that ideas can be more or less compatible with other ideas, you can almost picture how cohesive but mutually incompatible worldviews grow like crystals from a few divergent "seed" convictions. And it wasn't just the two of them. Segerstråle describes the formation of separate group belief systems on different sides of the conflict — belief systems that determined how various pieces of evidence was interpreted.

Thus, we can say that the two academic camps that had formed on the basis of the IQ and sociobiology controversies effectively came to live in two different worlds of factual knowledge, taken-for-granted assumptions, and attitudes towards such things as the media. Basic social psychological theory can make some predictions as to what will typically happen in a case of such pre-existing interpretive frameworks. Any incoming information will be accommodated in line with existing convictions; various well-known cognitive defense mechanisms will be operating to effectively protect the members of each camp from serious challenges to their existing 'knowledge'; and within each camp, members will reinforce one another's beliefs.

This in particular once convinced me that when you try to make sense of a disagreement you have to make the effort to understand the interpretive frameworks to which each side subscribes. It's difficult because it requires figuring out other people's underlying mental architectures that they themselves take for granted enough to be unwilling or even unable to articulate in communication. You have to look for clues and draw careful, tentative conclusions. It's like being a detective.

Planters, weeders, couplers and decouplers

Getting a glimpse what's in another's mind can be both fascinating and a little unnerving, like seeing the dark silhouette of a large animal under water. There were parts of the book where I was both fascinated and unsettled by getting a view of belief systems that at once made others' motivations easier to understand, but also felt threatening because of how incompatible with my own viewpoint they were.

For example, I'm a "planter" at heart. Segerstråle distinguishes between "planters and weeders in the garden of science" where planters focuses on science as the generation of new hypotheses, theories and explanatory models and "weeders" on criticizing the hypotheses/theories/models that aren't up to snuff. Members of the camp that agreed with Lewontin that not everything called science was created equal and that undercooked theorizing lent itself to political misuse and thus constituted a continuing threat of social harm, were more likely to be characterized as "weeders". Segerstråle quotes from "Not in our Genes" by Richard Lewontin, Steven Rose and Leon Kamin:

Critics of biological determinism are like members of a fire brigade, constantly being called out in the middle of the night to put out the latest conflagration, always responding to immediate emergencies, but never with the leisure to draw up plans for a truly fireproof building. Now it is IQ and race, now criminal genes, now the biological inferiority of women, now the genetic fixity of human nature. All of these deterministic fires need to be doused with the cold water of reason before the entire neighborhood is in flames. Critics of determinism, then, seem to be doomed to constant nay-saying, while readers, audiences and students react with impatience to the perpetual negativity.

The "planters" in question, who were more liberal with speculations, would in turn strongly disagree that their theories implied anything of the sort Lewontin and his coauthors claimed. Partly this reflects a difference of opinion on the relationship between science and society: planters tended to be scientific traditionalists who often believed that "is" didn't necessarily imply

"ought", that the freedom of inquiry was essential and itself a moral good and that restricting research output due to political concerns was a serious threat. Conversely, the weeders acted as if it was the norm for scientists to consider the social consequences of their research. Perhaps, suggests Segerstråle, they wanted to bring about such a norm by acting as if it already existed.

This might sound familiar to anyone who's read my article on the spat between Sam Harris and Ezra Klein, which shares a lot of characteristics with the sociobiology debate and ended with a discussion of "coupling" vs. "decoupling" attitudes towards science and society. Was it simply a conflict between "decouplers", who prefer to decouple scientific questions from their social context and implications, and "contextualizers" (or "couplers") who believe it's impossible to do so and dishonest to pretend you can?

You might easily think so, especially since my initial characterization of decoupling is in fact an amalgam of two ideas: Keith Stanovich's notion of cognitive decoupling — the practice of excluding the real world aspect of a scientific/logical problem and treating it as a formal exercise — and what Segerstråle calls "coupled thinking". She mentions it twice in the book, the first time as the belief that good scientific practice and progressive social views goes hand in hand and thus that correct science could virtually by definition not be used for politically regressive ends, and the second time more generally as the belief that those who disagree with one's own scientific views do so for political reasons.

It explains some of the controversy but by no means all of it. Some scientists on both sides did strive for separating science and politics, like Richard Dawkins, Peter Medawar, John Maynard Smith and Salvador Luria, but notably not Wilson himself. *Sociobiology* and his following books, especially his 1998 work *Consilience*, was part of a grand moral-scientific agenda that involved putting morality on a genetic foundation and remaking the humanities to be informed by biology.

Does that mean that Wilson was in fact every bit the conservative (or worse) hierarchy-hugger the critics imagined? Not really. While nature vs. nurture is often cast as conservative vs. progressive, this fight took place in academia and genuine conservatives were nowhere to be seen. To the extent that the controversy had a political dimension, it was a fight between radicals (often but not always explicitly Marxist) and liberals. Wilson's desire to work out a genetic basis for morality did clash with a certain radical conception of humanity, which made some critics code it as right-wing, but it, and he, was not.

In other words, he didn't quite play along with the villain role assigned to him by some critics. Segerstråle discusses the often highly uncharitable readings some of them had to subject his

words to in order to make him say what was required of him. His agenda was primarily *anti-religious*, not politically conservative. Wilson wanted to free people from the restrictive rules imposed on them by organized religion, and considered sociobiology a way to ground morality in something other than religious belief. In a way he was a bit like Lewontin in that he seems to have believed that true science (a "genetically correct moral code") would by default have beneficial social implications.

Dawkins and Gould: the next generation

This fighting of different, to each other orthogonal fights carries over from Wilson and Lewontin to their spiritual successors, Richard Dawkins and Stephen Jay Gould. Dawkins and Gould fought over the popular conception of evolution during the '80:s and '90:s, a rivalry Segerstråle discusses in the last third. To be very brief, Dawkins' view is centered on natural selection of individual genes and gradual adaptation to the environment as the core mechanism of evolutionary change, while Gould argues that this pure, algorithmic view is inadequate. He insisted that structural path-dependency and contingent events played large roles in evolution and couldn't just be brushed aside.

Just like the combination of political ideals and attitudes about good science animated Lewontin's criticism of Wilson, it seems as if, at least in part, Gould's distaste for the idea that society accurately rewards inherent merit animated his dislike of Dawkins's and others' narrow focus on natural selection and fitness at the expense of structural restrictions and the slings and arrows of outrageous fortune. It's hard not to hear loud political overtones in his tireless complaints about "adaptationism" — the tendency to think of biological structures as by default having evolved to serve a particular function by natural selection. Gould had a Marxist background like Lewontin and demonstrated a strong distaste for anything suggesting that evolution was somehow "fair" and that differential survival implied differential "fitness" on some objective scale.

Having been interested in their differences for a long time, my judgment is that their actual beliefs about evolution don't differ very much. There's even a case where Gould uses an analogy involving a cake (originally thought up by Patrick Bateson) to criticize Dawkins, one that Dawkins had in fact used to describe *his own* view in an earlier book. They disagreed about emphasis and interpretation, and that was most likely due to them being preoccupied with different issues outside biology itself.

Dawkins's enemy was, just like Wilson's, religion — more specifically creationism. His books were meant to show how blind evolution was capable of creating complex life of the kind we see

in the world. He intended to describe the mechanism that did the heavy lifting. Then there is lots of complexity and contingency involved in the actual, particular history of life on Earth, but Dawkins doesn't care as much about that because it's not load-bearing. It's part of what happened, sure, but doesn't do explanatory work for how it was *possible for it to happen without intelligent design*.

That stuff is all about incidentals, not essentials — noise rather than signal, and Dawkins doesn't mind putting it to one side because the details aren't what matters in the fight against creationism. However, if you are Stephen Jay Gould and want to focus on how in biology as in society, success is not determined by individual fitness or merit but is largely contingent and restricted by other factors out of individual control, then the incidentals, the "noise" part of the process, is exactly what you want to put front and center. Add to that a Lewontin-like view that good science requires real, exact details that Gould also seems to have shared, and I believe you have a decent account of that conflict as well.

Modern day descendants

The final third is both a retrospect of the sociobiology debate and the nature-vs-nurture debates that came with the increasing acceptance of biological explanations of human behavior (and the birth of evolutionary psychology) in the 1990s, and an exploration of just how many philosophical questions these issues manage to touch: the nature of humanity, free will, the relationship between different scholarly disciplines and the status, role and responsibility of science in society.

I came away with the impression that it's especially difficult to summarize a debate where so many people involved are top-tier intellectuals who no doubt have spent a lot of time and effort meticulously building, sculpting and honing their worldviews to be internally consistent, and hold sophisticated, individually differentiated views on everything. The chances of simplifying well are a lot worse than if we were dealing with holders of disconnected beans-in-a-bag opinions or standard-issue partisan packages.

I can't help but think that this is getting even more difficult today, for different reasons. There are many more voices than in decades past and extended dialogue or even just plain honest *communication* across ideological borders is the exception. The unrestricted flow of decontextualized bits of information on social media at the same time means that carefully thought up ideas of top academics regularly get picked up by us second-, third- and fourth-rate thinkers without the context needed to understand them fully. Given that, it gets harder to know just how much sense you can expect something to make. If Joe says "X!" and Jane says "No,

Y!", their exchange of rehearsed talking points might be playacting a conflict between whole schools of thought clustered around X and Y that they don't understand or even know about. And they might mistake each other's inability to justify X and Y to each other properly for proof of the inherent worthlessness of those beliefs. What is the best way to make sense of that disagreement? Do we go by person or by ideas invoked?

Circling back

This book really is as good as I remembered it, which is a huge relief. However, writing a positive review makes me a little bit suspicious with myself. Since 2008 I've learned to be more skeptical of that which caters to my own preconceptions and biases, so I can't help but wonder if I like it so much because it's relatively friendly to my own views. For example, Segerstråle is quite willing to call out the dishonesty with which some of Wilson's critics (most notably Stephen Chorover) twisted his words. She also paints a picture of Richard Lewontin's personality that isn't particularly flattering compared to her depiction of Wilson.

There's hardly anything in there about what she thinks herself. This is probably prudent, but it makes you wonder how fully you can trust her account. The only relevant info I could find was from [this article](#) in *The Chronicle of Higher Education*:

Although she admits that she thinks Mr. Wilson behaved better than his critics, Ms. Segerstrale says she can't say which scientist was right. 'I am not Wilsonian, and I am not Lewontinian,' she says, but in her book she tries to present balanced scientific criticism of both sides. As for the moral ambitions of the opponents, she sees them both as 'defensible'.

The same article quotes Irven DeVore, professor of biological anthropology at Harvard and minor character in the story, saying it's "unlikely that there will be another review of this tumultuous period in evolutionary biology that is so thoughtful and comprehensive." So I think my answer has to be that Segerstråle is likely as trustworthy on this as anyone can reasonably be. I would've liked to read some comments from Lewontin but according to the article he hasn't read the book and doesn't want to, stating that "I don't know what it'll tell me that I don't already know". In my opinion that doesn't speak in his favor.

Now after the second read I suspect I've been even more influenced by this book regarding how "erisology" is supposed to be done than I thought I was. It does a lot of work on its own to instill the kind of mindset I think you need in order to deal with disagreement well. It's full of quotable

passages and if you're at all interested in the science, politics and philosophy of human nature, I absolutely recommend it.

Democracy in America vol. I by Alexis de Tocqueville

Alexis de Tocqueville has a Frenchman who visited America in 1831 to report on how the American project and democracy work out in practice, and he published the first volume of his finding in 1835. It was an early work of political science and sociology. It's so densely packed with ideas that any chapter would easily fill a book if written today. Because it's so dense, and because of the academic standards of the time, it doesn't have the mathematical and historical rigor one would expect from a modern work. His analyses are determined largely by personal experience, intuition, and inspired by a handful of historical examples. But as a work of popular science, it compares favorably to something one might read today especially in one regard: he makes a number of specific predictions by which we can judge his models, and we've accumulated enough history by now to do so! Remind me to take another look at Turchin in 100 years. I've made a point to note all of Tocqueville's predictions and score him. However, it's hard to say how obvious these insights were at the time – many of his predictions were widely believed.

“Amongst the novel objects that attracted my attention during my stay in the United States, nothing struck me more forcibly than the general equality of conditions. I readily discovered the prodigious influence which this primary fact exercises on the whole course of society, by giving a certain direction to public opinion, and a certain tenor to the laws; by imparting new maxims to the governing powers, and peculiar habits to the governed. I speedily perceived that the influence of this fact extends far beyond the political character and the laws of the country, and that it has no less empire over civil society than over the Government; it creates opinions, engenders sentiments, suggests the ordinary practices of life, and modifies whatever it does not produce.”

As far back as history can recall, Europe had been ruled by hereditary aristocracies, mostly feudal monarchies. While Europe progressed toward democracy is slow, stuttering steps, the unique geographical and cultural initial conditions of America led to it being the most advanced democracy in the world. Tocqueville uses the term “democracy” almost interchangeably with freedom, equality, and liberalism, all of which seem to reinforce each other.

The Americans are descended principally from the English, and especially by the Puritans, who had a disproportionate impact on American culture. With vast amounts of arable land in America, and no landed gentry hoarding it, each man could generally provide for his own income. And as a rapidly expanding country, if some man were to acquire a great deal of land, it would often be better to move West than to pay rentiers. This produced an impressive level of equality, by 19th century French standards. In Europe, the rich and the politically powerful are almost synonymous; when most men are of comparable wealth and power, democracy is a natural form of government.

1830's America was a limited democracy in which voting rights were heavily restricted to varying degrees depending on what state one was in, but the franchise had been gradually expanding over the past decades, especially in the North. The state and federal governments were smaller and weaker than in Europe, or in 2021. American experience with democracy was dominated by local government in towns. Custom carried as much force as law, as both drew their authority from the will of the majority. Those living in the most Puritanical parts of the country were strictly bound by social norms, but they were voluntarily accepted, as it was easy enough to seek fortune elsewhere. While free, Americans are not generally independent. Rather, they organize in civil society to debate government affairs and organize interventions. Equality engenders freedom, which leads to travel, assimilation, commerce, and liberalism. Just about every aspect of American society and government follow from this structure.

Tocqueville lays out the three branches of government and the system of state and federal government. Tocqueville continually remarks on how little power the federal government and the president have, compared to European royalty, whereas the judiciary is stronger in America than in Europe. Here, it not only resolves disputes between individuals, but between arms of the government, which it can do by appealing to the constitution. Tocqueville saw a potentially dangerous exception to this rule: in the case of impeachment, Congress takes on the role of the judiciary. His next prediction:

“We must not be misled by the apparent mildness of the American legislation in all that relates to political jurisdiction. It is to be observed, in the first place, that in the United States the tribunal which passes sentence is composed of the same elements, and subject to the same influences, as the body which impeaches the offender, and that this uniformity gives an almost irresistible impulse to the vindictive passions of parties. If political judges in the United States cannot inflict such heavy penalties as those of Europe, there is the less chance of their acquitting a prisoner; and the conviction, if it is less formidable, is more certain. The principal object of the political tribunals of Europe is to punish the offender; the purpose of those in America is to deprive him of his authority. A political condemnation in the United States may, therefore, be looked upon as a preventive measure; and there is no reason for restricting the judges to the exact definitions of criminal law. Nothing can be more alarming than the excessive latitude with which political offences are described in the laws of America ... But I will venture to affirm that it is precisely their mildness which renders the American laws most formidable in this respect ... By preventing political tribunals from inflicting judicial punishments the Americans seem to have eluded the worst consequences of legislative tyranny, rather than tyranny itself ... When the American republics begin to degenerate it will be easy to verify the truth of this observation, by remarking whether the number of political impeachments augments.”

I'm giving half credit for this one. Impeachments are very rare, but many would argue they're sometimes politically motivated, and no credit for the claim that the republic will degrade due to ease of impeachment.

I can't help noting his thought on freedom of the press:

"The characteristics of the American journalist consist in an open and coarse appeal to the passions of the populace; and he habitually abandons the principles of political science to assail the characters of individuals, to track them into private life, and disclose all their weaknesses and errors.

Nothing can be more deplorable than this abuse of the powers of thought; I shall have occasion to point out hereafter the influence of the newspapers upon the taste and the morality of the American people, but my present subject exclusively concerns the political world ... The individuals who are already in the possession of a high station in the esteem of their fellow-citizens are afraid to write in the newspapers, and they are thus deprived of the most powerful instrument which they can use to excite the passions of the multitude to their own advantage."

If nothing else, this led me to reconsider whether there's been any significant change in the practice of journalism recently.

De Tocqueville is interested in the nature of American democracy in large part because it represents the future of Europe, which (due to initial conditions) is approaching democracy more slowly. Next prediction:

"And to no people can this inquiry be more vitally interesting than to the French nation, which is blindly driven onwards by a daily and irresistible impulse towards a state of things which may prove either despotic or republican, but which will assuredly be democratic."

And much later in the book:

"Those who hope to revive the monarchy of Henry IV or of Louis XIV, appear to me to be afflicted with mental blindness; and when I consider the present condition of several European nations—a condition to which all the others tend—I am led to believe that they will soon be left with no other alternative than democratic liberty, or the tyranny of the Caesars."

Mostly correct. France would soon enter the Second Republic and soon the second Empire, as Napoleon was elected and later emperored. And in the long run, France became a first world democracy.

While the American government is smaller (per capita) than European contemporaries,

"I have no hesitation in predicting that, if the people of the United States is ever involved in serious difficulties, its taxation will speedily be increased to the rate of that which prevails in the greater part of the aristocracies and the monarchies of Europe."

Correct, this happened only decades after publication.

One of the most obvious problems with democracy, even at the time the constitution was drafted, was tyranny of the majority. In this free society, the majority may have even more power than in France, as the majority has the government behind it, and the government has more moral authority when it represents the people as a whole.

“In America the majority raises very formidable barriers to the liberty of opinion: within these barriers an author may write whatever he pleases, but he will repent it if he ever step beyond them ... If ever the free institutions of America are destroyed, that event may be attributed to the unlimited authority of the majority, which may at some future time urge the minorities to desperation, and oblige them to have recourse to physical force.”

Fact check: free institutions not destroyed yet.

The last major section of Vol I addresses the questions of race which presented obvious problems for American society. Tocqueville is pessimistic regarding the fate of Native Americans. He almost never cites statistical evidence, such as estimates of the Native American population, relying entirely on his personal observation. But, those are quite compelling.

“A few European families, settled in different situations at a considerable distance from each other, soon drive away the wild animals which remain between their places of abode. The Indians, who had previously lived in a sort of abundance, then find it difficult to subsist, and still more difficult to procure the articles of barter which they stand in need of ... At length they are compelled to acquiesce, and to depart ... it is insinuated that they have not the means of refusing their required consent, and that the government itself will not long have the power of protecting them in their rights ... Half convinced, and half compelled, they go to inhabit new deserts, where the importunate whites will not let them remain ten years in tranquillity ... I believe that the Indian nations of North America are doomed to perish.”

The Native American population continued to decline after Democracy in America was published. In New England, they were almost completely displaced. But, the Native population leveled off, and has been growing as a fraction of the population since circa 1950.

“Since the first settlement of the British colonies, the number of inhabitants has about doubled every twenty-two years. I perceive no causes which are likely to check this progressive increase of the Anglo-American population for the next hundred years; and before that space of time has elapsed, I believe that the territories and dependencies of the United States will be covered by more than 100,000,000 of inhabitants, and divided into forty States.”

Correct! In 1935 there were 48 states with a total population of approximately 127 million.

“But the case of Texas is still more striking: the State of Texas is a part of Mexico, and lies upon the frontier between that country and the United States. In the course of the last few years the Anglo-Americans have penetrated into this province, which is still thinly peopled; they purchase land, they produce the commodities of the country, and supplant the original population. It may easily be foreseen that if Mexico takes no steps to check this change, the province of Texas will very shortly cease to belong to that government.”

100% correct.

De Tocqueville advocates for no particular policies in the book, only making observations and predictions, but his preferences are often clear. He describes slavery as an evil, harmful to Black people, and Whites, and society generally. Historians sometimes argue that slavery diverted resources away from industrialization in the South, contributing to its loss in the Civil War. Tocqueville goes further, and argues that the institution of slavery caused any kind of work at all to be associated with Blackness, and discouraged European Americans from laboring any more than necessary, or even desiring the wealth that labor might provide. At the time of the writing of *Democracy in America*, the North was in the process of eliminating slavery, state by state. But whenever a Northern state is about to do so, its slaveowners immediately sell their slaves to the south before they lose value. As such, the number of slaves in the US stays almost constant, but they're increasingly concentrated in the South. As a result, the Southern slave states have an ever increasing investment in the institution of slavery, and more to fear should it be overturned. This incremental advance of abolition could never convert the entire country, as the whites in the last remaining slave holdouts would have too much to lose. With national and global mood turning towards abolition, this was clearly unsustainable, and Tocqueville thought the result would be a civil war; but he imagined it as a slave revolution, a war between the Southern Blacks and Whites. Tocqueville thought that a civil war between states was very unlikely, as each had too much power of self governance for the Union to stop them from seceding.

“By the choice of the master, or by the will of the slave, it will cease; and in either case great calamities may be expected to ensue. If liberty be refused to the negroes of the South, they will in the end seize it for themselves by force; if it be given, they will abuse it ere long.”

75% credit. The Civil War happened not too long after publication.

“The time will therefore come when one hundred and fifty millions of men will be living in North America, equal in condition, the progeny of one race, owing their origin to the same cause, and preserving the same civilization, the same language, the same

religion, the same habits, the same manners, and imbued with the same opinions, propagated under the same forms ... The Americans of the United States must inevitably become one of the greatest nations in the world; their offset will cover almost the whole of North America; the continent which they inhabit is their dominion, and it cannot escape them. What urges them to take possession of it so soon? Riches, power, and renown cannot fail to be theirs at some future time”

True.

“I am strangely mistaken if the Federal Government of the United States be not constantly losing strength, retiring gradually from public affairs, and narrowing its circle of action more and more. It is naturally feeble, but it now abandons even its pretensions to strength. On the other hand, I thought that I remarked a more lively sense of independence, and a more decided attachment to provincial government in the States. The Union is to subsist, but to subsist as a shadow; it is to be strong in certain cases, and weak in all others; in time of warfare, it is to be able to concentrate all the forces of the nation and all the resources of the country in its hands; and in time of peace its existence is to be scarcely perceptible ... the Government of the Union will grow weaker and weaker every day.”

This is completely wrong; the federal government has been gradually growing in power relative to the states. We've been fighting wars most of the time since WWI, maybe that did it.

“There are, at the present time, two great nations in the world which seem to tend towards the same end, although they started from different points: I allude to the Russians and the Americans. Both of them have grown up unnoticed; and whilst the attention of mankind was directed elsewhere, they have suddenly assumed a most prominent place amongst the nations; and the world learned their existence and their greatness at almost the same time ... Their starting-point is different, and their courses are not the same; yet each of them seems to be marked out by the will of Heaven to sway the destinies of half the globe.”

Not bad!

“... the Spaniards and the Anglo-Americans are, properly speaking, the only two races which divide the possession of the New World. The limits of separation between them have been settled by a treaty; but although the conditions of that treaty are exceedingly favorable to the Anglo-Americans, I do not doubt that they will shortly infringe this arrangement. Vast provinces, extending beyond the frontiers of the Union towards Mexico, are still destitute of inhabitants. The natives of the United States will forestall the rightful occupants of these solitary regions. They will take possession of the soil, and establish social institutions, so that when the legal owner arrives at length, he will find the wilderness under cultivation, and strangers quietly settled in the midst of his

inheritance ... The province of Texas is still part of the Mexican dominions, but it will soon contain no Mexicans; the same thing has occurred whenever the Anglo-Americans have come into contact with populations of a different origin.”

Right again.

I've skipped over a lot due to length. Tocqueville addresses the natural geography of North America; the maritime trade; religion; details of French and British society; and more. In Vol II, which I haven't addressed, he discusses philosophy among the Americans, the arts, the situation of women, capitalism, the military, and more.

I'm left with a few impressions. One is that despite Tocqueville's protests, the book is suffused with historical determinism; the fate of peoples was marked out long ago by their initial conditions. He opposes this view on the grounds that it's everyone's chosen actions that create these trends. And yet he continually makes predictions about what, precisely, everyone will choose. In some cases, he makes his case very strongly, his predictions have borne out, and we can even see them borne out repeatedly as more countries democratize.

Another is that his predictions are sometimes wrong. Social scientists today seem to have just as much difficulty despite what appear to be huge advances in theory and data collection. Every historical pattern has exceptions, all patterns are debated, and making reliable predictions seems impossible. It's unclear whether any level of intelligence and rigor can solve this problem.

Is there value in studying history and social science beyond academic curiosity? If study of these fields can make reliable predictions, that would suggest historical determinism is true; and if determinism is true, studying social trends will make no difference. Maybe there's a historical trend that they're useful to people who study them, but not in a way that affects society at scale? Or maybe we should just be happy with academic curiosity.

Disunited Nations (2020) vs Dawn of Eurasia (2017)

What does the future look like? We are living through a transition between epochs. Whether marked by COVID-19, the election of Donald Trump or earlier by the global sub-prime crisis, the golden age of post-Cold War prosperity is ending. With the era defined by US political, cultural and economic hegemony, its decline is inextricably linked to the decline in US influence. Is the twenty first century really going to be the “Asian century” as China’s growth continues unabated? Or perhaps African, given by far the largest forecast population growth? What will become of the US? Of China and Russia and Europe? Two thinkers have sought to define this future.

I first came across Bruno Maçães in 2017 on [Marginal Revolution](#) where Tyler Cowen was effusive about Maçães’ new book. I have enjoyed following his conversations and [thoughts](#) ever since, but it was only recently that I read *Dawn of Eurasia*. It is the first book of a career politician and diplomat clearly in love with his continent.

Peter Zeihan I came across on Patrick O’Shaughnessy’s excellent [podcast](#). His brash prophesy and contrarian views on geopolitics are hypnotic and endlessly fascinating. *Disunited Nations* is his latest in a series that documents the rise and rise of US power. I found comparing them irresistible. Each lingers after reading. It’s that wonderful feeling of discovering a new area of knowledge to mine. Not natural companions, and mesmerising in their own ways, each story has a different texture and plots a different path for the world. Where one sees pessimistic reversion to a historic state of conflict, the other sees hopeful evolution. Where one deterministically condemns nations to their geographic destinies, the other sees each nation’s destiny as unwritten, yet to be informed by its history, literature and peoples.

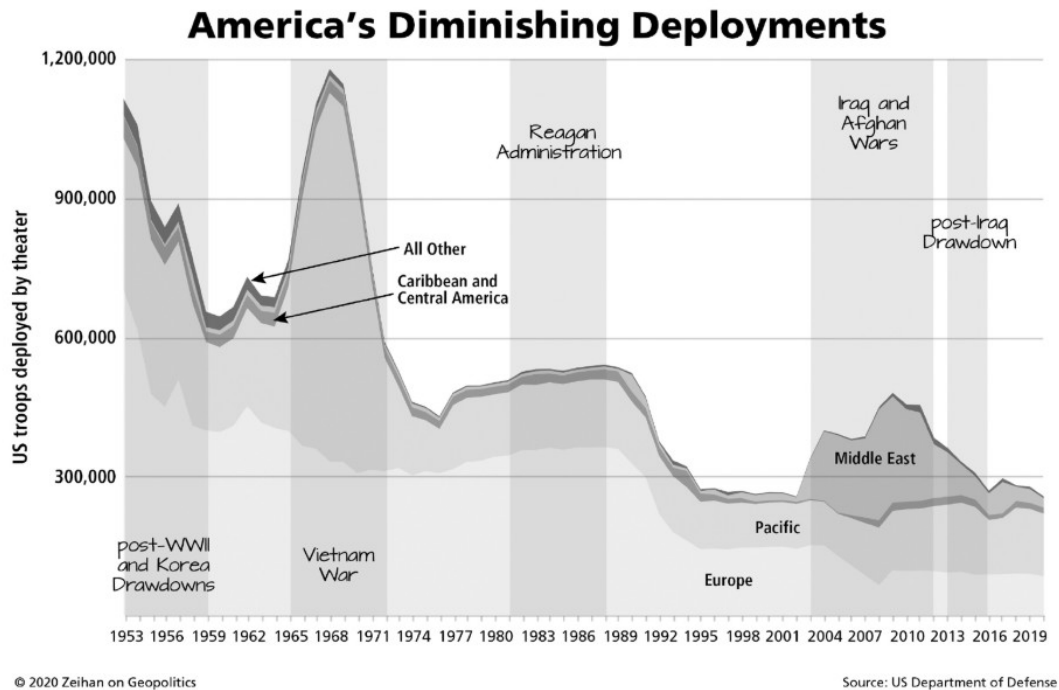
Both Peter Zeihan and Bruno Maçães see US influence receding. But they agree on little else. Zeihan is deeply pessimistic about a world that awaits a more isolationist US, with a crumbling world order leaving less room for prosperity and reverting to nation-states jostling for food, energy and military security. Maçães sees China’s rise as a

harbinger of consolidation on the Eurasian supercontinent, with Europe, Russia and China becoming increasingly interlinked politically and economically. China must grow towards Europe as the US withdraws.

Their most obvious conflict is that Zeihan rejects a “dawn of Eurasia”. Where Maçães foretells super-continental consolidation, Zeihan expects devolvement into tense state competition. But these kinds of differences are less interesting than the differences in their underlying approaches: can there ever be a unifying theory of the future? Can we march forward, to a new place, or is civilisation destined to repeat itself? Can we build our future or are we pulled along the pre-cut canals of the world we inherited?

The US World Order is coming to an end

Zeihan’s *Disunited Nations* seeks to answer one question: what happens to the prospects of nations when the US withdraws from world affairs? What happens when the post-WWII World Order – along with all the networks of alliances and protected trade routes that comprise it – disintegrates? And the US *will* withdraw, according to Zeihan. In fact, by global troop deployment, it’s been withdrawing for a while:



The two things that kept the US interested in maintaining the world order have disappeared: an arch enemy in the Soviet Union and reliance on foreign energy. Continuation of the Order since the collapse of the USSR over the last thirty years has largely been a result of geopolitical inertia. And energy independence arose with the advent of fracking, when the US became a net crude-oil and products exporter as of 2020.

How exactly does the US maintain the Order? Two ways: the US Navy and free trade.

The Cold War led the US to buy allies into its orbit, through global subsidies via freer trade, protected by naval supremacy. Free trade was never the end, but rather an effective tool to enrich US allies, to buttress against Soviet expansion, says Zeihan. Take Japan, the obliterated former enemy of the US and Axis power: its average Japanese citizen became wealthier than the average US citizen in the mid-eighties. The US effectively subsidised the Japanese citizen to greater wealth than its own. China today is credibly touted as the next global superpower. But it could only become the factory of the world because its shipping routes are protected by US naval supremacy.

Post-Cold War, the US no longer has a dog in the fight. It's increasingly looking for a way out, especially after its two decades long escapades in Iraq and Afghanistan. So what now? It's... pretty bleak:

“Without the global security the Americans guaranteed, global trade and global energy flows cannot continue. Seven decades’ worth of global industrialization and modernization are not simply at risk, the very pillars of civilization are cracking. In a world without stability, the questions become: Who was most dependent upon the world that was and so will fall? And who was most restrained by the old Order and so will soar?

The world we know is collapsing. Entire countries are watching in horror as what makes them possible—global access, imported energy, foreign markets, American troops—slips through their fingers. For many, there just isn't enough access or energy or markets or security for them to maintain what they have, much less grow. In a world of want, the questions become: What do countries need to survive in a scrambled world? Who will shoot to get what they need? And who gets shot at?

Not all competitions and scarcities are created equal. Nearly all food is dependent upon global trade, whether in the form of imported inputs or the foodstuffs themselves. For decades, the world's experiences with famine have been crises of distribution, the inability to match foods with mouths. Global breakdown guts food supply itself. The security concerns of the past two decades were largely limited to terrorism, but the tools necessary to counter terror are radically different from In a world of different scarcities and different tools, the questions become: Where will trade patterns hold and where will they collapse? Which ones are worth fighting over? Which tools will be brought to bear? Are we on the verge of a mess of overlapping and interlocking naval competitions for something as basic as the right to eat?”

Zeihan's post-US order is one of Hobbesian decline. Whilst US power persists, its position is characterised along similar lines to those used by Maçães: *“a splendid*

isolation, born of power and promising many degrees of freedom". One of the biggest losers, on the other hand, is China:

"The question is not whether China can be the next global hegemon. It cannot. The real question is whether China can even hold itself together as a country."

Given the current zeitgeist around China rising, a view to which Maçães subscribes, it is Zeihan's most contrarian view in a book full of contrarian deviations. Instead of China or Germany or Russia rising, he sees this century as dominated by the US, Japan, Argentina, France and Turkey.

The US Navy patrols the world

Following WWII, the US basically had the only navy left. And with the rest of the world in tatters, it had its pick of global naval base positions. This network today is impossible to replicate. It comes with strategic base positions, decades of naval investment and technological advancement, and operational expertise gained through combat. And Zeihan makes the point with typical panache:

"...aside from the Americans, no one floats even a single fully functional supercarrier, much less a supercarrier battle group, much less a global naval force. The next-largest non-American carriers aren't even really operational. Russia's Kuznetsov appears to like catching on fire; the UK's Queen Elizabeth hasn't yet exited sea trials; India's Vikramaditya is a floating mass of technical problems; and the engines on France's Charles de Gaulle aren't powerful enough to bring it up to aircraft-launching speed except under near-perfect conditions. In fact, the next nine largest fully operational carriers are also all operated by the US Navy.

But Zeihan also makes a deeper point than the happenstance of US naval supremacy. *Of course* the US Navy was *always* going to be a naval power. The US is protected by seas and is the rogue child of the greatest naval nation in history, Britain, which is itself an island. And, *of course*, Britain would have had the greatest navy in history, because it didn't have Germans and French and Prussians and Russians and Mongols marauding across its borders. This is the heart of Zeihan's worldview. He is a geographic determinist. Show me a nation's geography, he says, its weather, its topology, and I will show you its capacity to feed itself, to build infrastructure, to hound its neighbours for energy or be hounded in turn. The US has all it needs where it sits and, by virtue of its surrounding oceans, it's very hard to invade, and unlikely to disintegrate given the contiguous landscape within. Saudi Arabia, for example, has fewer choices, being a land of sand and oil. China has for eternity battled to maintain its internal cohesion, while being constantly threatened by hostile neighbours on all sides. It could never build a real naval capability by virtue of how disparate its threats are, spreading its focus thin across strategic imperative. In fact, Zeihan can't stress the meagre nature of China's current navy and future naval prospects enough:

“With its naval order of battle limited by its small number of long-range ships, China is horrifically dependent upon the ports of others. China's only reliable ally—stretching the meaning of reliable and ally—is North Korea. As China borders North Korea, North Korean ports—not positioned or designed to support large-scale naval power projection—don't exactly help the picture. China has therefore pushed a “string of pearls” policy to create a line of friendly ports between the Chinese coast and areas of interest. Malaysia, Myanmar, Sri Lanka, Pakistan, and Kenya, for example, serve as Chinese pearls on the all-important energy-shipping route from the Persian Gulf. A big piece of China's One Belt, One Road (OBOR) diplomatic effort is to rope such countries into China's sphere of influence. The path the Chinese have followed is rather sophomoric: Beijing insists that the funds it has provided to these countries were not grants, but instead loans that need to be paid back. At the time of this writing, relations with all five countries have soured

expressly because of OBOR to the point that China cannot trust any of them to serve as ports in any sort of military storm. (And that's before one considers the general hostility of their neighbors—most notably India—to China mucking about in their neighborhood.) That leaves the Chinese with only one fully fledged foreign base: Djibouti. It is a facility over five thousand miles from China's shores, designed to serve as an operational node to help multinational anti-pirate operations. It isn't so much that the Chinese navy's only overseas base exists to hunt half-starved black dudes in speedboats. The point is that China had to go so far from home to find a country poor enough with few enough preconceived notions about the Chinese that it would be willing to host a full Chinese base in the first place. Not to mention that ironically—hilariously—China's Djibouti base is possible only because the Americans perceive it as helping uphold the lagging Order, and because Djibouti is so popular for anti-piracy bases that the Chinese hardly enjoy a strategic monopoly there: France, the United States, Italy, Saudi Arabia, and the United Arab Emirates all operate permanent facilities in Djibouti.”

Well, what about its supercarriers?

“Much is made of China's Liaoning . It shouldn't be. On paper the Liaoning has perhaps one-seventh the combat capacity of a Nimitz [US's first modern supercarrier]. . . and the Chinese didn't build it. The Soviets did. In the 1980s. Sort of. This first-of-type vessel was never completed. After the Soviet breakup, the Ukrainians stripped it for parts. The ship then languished, rusting at dock for a decade until the Chinese purchased it. Today it is used purely as a training platform. China's naval engineering is improving—the 001A , China's domestically manufactured clone of the Liaoning, is undergoing sea trials at the time of this writing, but there are still decades of work to go before China can float something that can compete with a Nimitz . Carrier engineering isn't rocket science. It's a lot harder.”

But again, more important than merely being behind, China is constrained by its inherent geographical position:

“China is an inveterate land power that has fought major land wars with each and every one of the powers it borders. It simply cannot afford the sort of resource focus that made the British navy possible.”

Maçães on the other hand does at one point consider a US conflict with China at sea:

“...if the next few decades witness a naval conflict between China and the United States that conflict will more likely be centred in the Indian Ocean than the Pacific, thanks to its greater strategic importance, and in that case India and the Indian navy will be the decisive factor.”

Maçães seeks to better ask the question: where might a conflict take place, and what might swing the balance? For Zeihan, there is no question: US naval supremacy leaves no room for one. One of them seems to be wrong.

How important are carriers and today's navies anyway? In his book *How to Defend Australia* Hugh White, a leading Australian defence strategist, doubts whether the navies Zeihan relies upon will even really matter in a major conflict:

“Warships will remain valuable for operations in waters that are not contested by other maritime powers, and likewise carriers and amphibious ships will remain useful in uncontested waters against less capable adversaries. But their roles in major maritime conflicts will disappear. Instead, war at sea will be dominated by submarines, aircraft, drones, missiles and satellites.”

Technology changes. Could navies be largely obsolete already? White is frank about the limits our knowledge today:

“...after many decades without major air or naval battles, we have little idea of how modern high-intensity combat will work, and which performance factors will make the most difference. For example, although huge sums have been spent on improving the stealth performance of fighter aircraft, no one really knows what difference stealth will make to the next great air-combat campaign. It might prove to be decisive, or to have been a waste of money.”

Maybe White is hedging his bets and Zeihan is simply calling it. Maybe Zeihan knows something White does not. But technology advances and wars surprise. This is the challenges of Zeihan's approach, which I'll return to further below. But first: how Zeihan's sausage is made.

Zeihan's geographic determinism

Zeihan's geographic determinism has a formula. Under a chapter 'How To Be A Successful Country', which uses China as its key counterexample, Zeihan outlines the four keys to success:

- 1. “Viable home territories, with usable lands and defensible borders*
- 2. A reliable food supply*
- 3. A sustainable population structure*
- 4. Access to a stable mix of energy inputs to participate in modern life”*

Grind a country through these layers and you will see how they will fare out in the wild when the US leaves its post. Take food, which under the Order has seen a *“fourfold expansion in food cultivation since 1946 and thus also the tripling of the global population.”* Per Zeihan, this has been possible because under the Order countries can import food from abroad or import the inputs required for food production to produce food themselves. Without this, *“a billion people are going to starve”*.

Again, Zeihan sees the point compounded in the case of China (it's a long quote but worth it):

“China faces a quadruple bind:

- 1. China's margin of error starts razor thin, and not only because its lands are below subpar. One downside of China's massive population is that the country has less farmland per person than Saudi Arabia.*
- 2. As China's population urbanized under the Order, much of the country's good(ish) farmland was paved over, pushing Chinese farmers farther inland into ever-more marginal territories, which require more and more inputs to produce the same amount of foodstuffs. Unsurprisingly, the Chinese economic sector that is most dependent upon the expansion-at-all-costs financial strategy is agriculture. When the Chinese financial system cracks, not only will China face a subprime-style crisis in every economic sector simultaneously, but it will also face famine—even if nothing goes wrong externally.*
- 3. On the surface, it appears China has sufficient oil and natural gas production to maintain domestic production of its fertilizer and fuel needs for its agricultural sector. After all, while China is a net importer of both fossil fuels, unlike most of Europe it still retains significant local production capacity. Not so fast. In any sort of constrained import environment—such as problems in the explosion-heavy Middle East—the Chinese will have to choose what they will let go of. Electricity? Motor fuels? Fertilizers? There won't be enough to go around, and that forces choices.*
- 4. China now isn't simply the world's largest importer of rice, barley, dairy, beef, pork, fresh berries, and frozen fish by tonnage. It absorbs more globally traded sorghum, flax, and soy than the rest of the world combined. The ongoing import of those products requires both the American Order and the ability of the wider world to produce the products in the first place. Unfortunately, it isn't as if the Chinese are alone in this dilemma. Roughly one-fifth of the world's agricultural produce is internationally traded, while four-fifths is dependent upon externally supplied petroleum-derived inputs, mainly fuel and pesticides. Remove the*

stability and shipping options and supply chains the Order makes possible, and many, many countries must figure out how to feed their populations with limited outside assistance. Most will fail.”

Like I said: bleak.

Maçães assumes the rise of China from the perspective of its own dreams and aspirations. From the Belt and Road policy objectives, its historic parallels to the Silk Road that once bound Eurasia, and the perspectives of its political and academic classes, China's assertiveness feels cautious yet determined. Maçães treats China's threat to the US seriously:

“In the same way that China and Japan had to change and adapt to the arrival of European civilization in the nineteenth century, the United States is feeling the impact of China's unique takes on modernity and capitalism”.

“...the United States has much to lose in its relationship with China, where differences in the respective political concepts and increasing parity in economic power pose distinct threats. But then there is also much to gain, as shifts in position from Beijing can help the United States solve many of its security and economic challenges.”

This seems in line with predominant view of China in the West today, which is viewed with a mixture of suspicion and awe. It is in direct conflict with Zeihan's insistence that the Chinese threat is wildly overblown. While Maçães elicits the wonderful texture of a nation through a local eye and a diplomat's intimacy with political process, some of his allusions to grander arcs of history such as these appear presumptive. His objective isn't quite as narrow as Zeihan's geopolitical forecasting, of course. He doesn't put forward a comparable analytical model, and the multi-lensed approach he takes, with its cultural and political appreciation, is in some ways the point. But Maçães does stake out geopolitical claims: consolidation in Eurasia, the corresponding threat to the US. Is

Zeihan a contrarian provocateur or does Maçães harbour uninterrogated establishment biases?

The Dawn of Eurasia

The Dawn of Eurasia is a gorgeous insight into the peoples and places of Eurasia, the supercontinent containing Asia and Europe. Maçães' perspectives on the whole arise from a focus on the smallest of its parts. It is in places like Chechnya, Azerbaijan, and Kazakhstan, between the tectonic plates of great powers, that new landscapes arise and hybrids form. These places are both a source of intercontinental tension and gaps where peoples and countries have fallen out of memory. The mystical lure in his note on the Caucuses might also hold true of the others: *"the past is part of the present. The successive waves of new peoples, religions and beliefs created a human landscape resembling geological sediments..."* It is in these layers and flows of people that he seeks to show us the future. The nature of Chinese dreams and the Russian spirit, Indian abstraction vs Western thought, the expansion of Europe across the world in a way that eluded Christendom. How these stories infuse the national ambitions of their peoples.

"Think of all the important and still undecided international questions of the last ten years. Energy security. Islamic radicalism. Ukraine. The future of Turkey and its position in the global system of alliances. The refugee crisis. They all point to the borderlands dividing Europe and Asia and are a direct result of flows – of people, goods, energy and knowledge – made possible by the gradual decline or collapse of the barriers keeping the two continents apart."

Maçães' journey starts in the Caucuses, which provides the first contrast to Zeihan: that of style as well as of character. Zeihan speaks of the Caucuses offhandedly as some beautiful but backward land of mountainous barbarians: *"If it weren't for the locals' tendency to flay outsiders, I'd go backpacking there."*

Maçães, on the other hand, delights in the layers of peoples and languages and cultures – from the last Jewish village outside of Israel, to the gaudy marble city of Avaza, Turkmenistan. Zeihan’s backpacking evokes a loud (and occasionally sneering) American lambasting through foreign lands. (My time in the rugged mountains of northern Georgia was far more weighted to eating *xhinkali* (giant dumplings) than flaying). Maybe that’s harsh on Zeihan, maybe he wasn’t referring to Georgia, but instead to Dagestan or the Chechnyans? Well, Maçães has an entire chapter about his time in Grozny, where the Chechnyan Muslims live under Russian auspices, facing Europe. He leads us through the Azerbaijani borderlands, where the local Muslims are proud custodians of the last Jewish village outside of Israel, while maintaining hostilities with their Christian Armenian neighbours. He contemplates the role of contemporary art in Iran. He admires the quickly growing and evolving trading hubs on the Khazak / China border, that meld into a flow of goods and people, whether Arab, Chinese and Khazak. The Caspian Sea’s *“energy resources are of great interest to Europe, China and Russia... it brings five coastal countries inextricably together, creates numberless variables of interdependency and forces them to cooperate, while at the same time bringing competition to a high pitch...”* These crossroads, the revival and pollination across the ancient Silk Road is entirely the point. As Maçães is seduced by the ebb of peoples and histories, so he conjures a future where these ancient cultural and geopolitical titans naturally converge. The direction of history that united Europe into the European Union continues under the Eurasian Economic Union and China’s Belt & Road initiative. Eurasia rises.

There is a geographic determinism at play here also. Why *shouldn’t* the largest supercontinent with the most advanced civilisations on earth finally unite? Maçães does not directly contemplate this. China is indomitable. Europe is not going anywhere. Russia is underrated. But perhaps more than some inevitable pull in their geopolitical futures is the sense that anything is possible, that the East has come so far and their will is so strong, that they will be able to make their futures.

Maçães is not only more hopeful about China. He approvingly cites a quote in the Financial Times about Britain: *“does Brexit offer Britain new opportunities as an agile trading nation, ‘a giant Atlantic Singapore’? Is a Eurasian capital being born on the shores of the Thames?”* Whereas Zeihan believes the British will be *“supplicants with no other options, so the terms of American-British interactions will be wholly American-determined.”* How will Britain be relegated to US vassal status? The end of the Order and Brexit condemn *“the Brits to a multiyear depression”*:

“The weird makeup of the early-Disorder British economy— minimal European access, but large-scale North American access—makes the British dependent upon the Americans for growth. With the American population aging far more gracefully than the British, this dependency is for the long haul and will become more intense with every passing year.”

Where Zeihan’s future is *“unavoidable”*, Maçães’ is yet to be born, hopeful.

Maçães’ is an altogether poetic vision, enthralled as he is by his subjects. His world is ancient or futuristic, lost cities and lost peoples (Circassians and Khazars), forgotten cities (Astrakhan), cities remade into the gaudy image of men, cities of the future, cities full of spies and lovers and adventurer archaeologists and diplomats and wily traders. Maçães is part Odysseus, enchanting us with his own travels and the tales that surround them. On the Caspian:

“Stepan Razin, that impossible combination of Cossack and a Caribbean pirate, who set out for adventure and plunder on the Caspian, endlessly tormenting the Russians along the northern shore and Persians in the south until his gruesome execution in the Red Square in Moscow in 1671.”

Or when he recalls Neft Dashlari, *“the first of the Caspian’s imaginary cities”*, a *“full city at sea with hundreds of kilometres of roads built on piles of landfill connecting different*

oil platforms, partially submerged apartment buildings hosting thousands of oil workers, schools and cinemas, hotels and even a tree-lined park.”

Maçães’ travels lead to pithy insights such as: *“modern culture so obviously shares the rootlessness that distinguished our distant nomadic ancestors,”* gathered from contemplating the nature of modern Kazakh life.

Maçães is also entranced with the Eurasian landmass itself. The way lands occupy our minds may be a function of the shapes and textures of the lands themselves. Consider how Russia’s eastern empires have never been considered alien to Russia in the way that Britain’s had. Its contiguous land mass has cast them as a natural appendage. Similarly, Maçães refers to Halford Mackinder, who suggested that *“the reason we never thought of Asian and Europe as a single continent is that seamen could not make the voyage around it.”*

Maçães draws a nation’s dreams and conceptions of itself from its novels and sci-fi, from its local fashionistas and traders. He footnotes, lots. He is in the detail, seeking to glean the future from this or that political speech or infrastructure project. Maçães is interested in institutions - their political endowments and governance mechanisms. Whether political and economic power reside in the Eurasian Economic Union vs the European Union. Considering the aspirations of the European Union to automate decision-making, for example, comes naturally to a former Secretary of State for European Affairs of Portugal who would have sat in on those governmental processes. Zeihan doesn’t really footnote. Zeihan shows us his four-point toolkit, then applies it country after country. His prophesy reads like a game of Risk, with nations playing against each other on a board, their strategic virtues and vices tallied up neatly for review. There is no room here for a nation’s conception of itself, of seizing its own destiny, of its poetry and lust, no anecdotes of pirates and princesses. These are all details, implicitly subsumed in the tectonic incentives that will necessarily propel countries in predetermined ways.

Maybe Maçães is too much of a romantic, seduced by the possibilities beyond the steppes. Maybe Zeihan is too reductionist. A Star Wars vs Star Trek-like competition in visions.

Determinism vs Hope

Where Zeihan's future is determined, Maçães' world is one of open slather. They sit across each other over the determinate vs indeterminate conception of history. They are not alone. Maçães has Peter Thiel and David Deutsch as companions, who virulently contend that the future is for us to create. If Thiel yearns for an [optimistic determinate](#) society, Zeihan is a pessimistic determinate, as he foresees an inevitable decline in world affairs. Maçães is probably an optimistic indeterminate, if optimistic mainly by temperament.

Thiel and Deutsch stand against theorists like Jarred Diamond, who in *Guns, Germs and Steel* argued that Western civilisational dominance stems from the fruits of its geography – the landscape, domesticable animals, climates. These allowed some peoples to conquer seas and far off shores, while constraining others from the ability to build cities, grow populations, scale armaments and win. Zeihan's thesis is a geopolitical extension of Diamond's geographic deterministic thesis.

Deutsch, critiquing Diamond's thesis in *The Beginning of Infinity*, contends that “[t]he conditions for a beginning of infinity exist in almost every human habitation on Earth.” This aligns neatly with Maçães. Maçães' dwelling on the deeper questions of cultural motivation and possibility is connected to the sense of agency his hope implies. Deutsch considered that “*mechanical interpretations of human affairs not only lack explanatory power, they are morally wrong as well, for in effect they deny the humanity of the participants, casting them and their ideas merely as side effects of the landscape.*” That is why humans are entirely absent from Zeihan's conceptions of state conduct, and at the heart of Maçães'.

How much of the future can anyone ever see?

In some ways, Zeihan is proposing his own theory of everything. Through his prism, he purports to forecast everything from a multiyear depression in Britain to “*a sub-prime style crisis in every economic sector simultaneously*” and famine at the same time in China. He is a little fast a loose with timing, and so perhaps the magic is in that elusion. (To be fair, on a [recent podcast](#), Zeihan did call China’s demise for this decade). His forecasts regarding Britain post-Brexit is off, for now at least – the UK economy is doing remarkably well, despite Brexit doomsayers. Regardless, his theory of everything is a little suspect, stretching powers of prophecy across geography, demographics, finance, military strategy, global supply chains, infrastructure costs. If he were really that powerful, Zeihan would perhaps be more at home in a hedge fund (for all I know, he is). Or perhaps his powers are more relevant over a fifty year view, in which case, it’s just hard to know.

Howard Marks has [written](#) extensively on the possibility of forecasting:

“So forecasting is difficult for a large number of reasons, including our limited understanding of the processes that will produce the future, their imprecise nature, the lack of historical precedent, the unpredictability of people’s behavior and the role of randomness, and these difficulties are exacerbated by today’s unusual circumstances.”

This was written in relation to COVID-19 and financial markets and economies, but the same could be said for world events. Are these times unprecedented? Inarguably, with the state of technology, rapid Chinese growth and urbanisation, and sheer US military might. Both Zeihan’s and Maçães’ theses propose as much.

Marks later quotes John Kenneth Galbraith:

“We have two classes of forecasters: Those who don’t know – and those who don’t know they don’t know.”

Zeihan comes out of a career at Stratfor, which calls itself “*an American geopolitical intelligence platform*”. On Stratfor’s founder, Wikipedia notes that “[George] Friedman’s reputation as a forecaster of geopolitical events led *The New York Times* magazine to comment, in a profile, “There is a temptation, when you are around George Friedman, to treat him like a Magic 8-Ball.”” Can geopolitical forecasting be a function of style as much as of analytical skill? Well it certainly makes for a riveting story, and as others have pointed out, a [story](#) goes a long way, even if it is cause for [suspicion](#).

Have there been historic attempts at this kind of prophesy? Diamond claims no powers of forecasting. This kind of forecasting would seem too ill defined and beyond the remit of superforecasters in the Philip Tetlock mould. How many hits will Zeihan have, and how much of it will be chance over prescience? Could there not have been a Zeihan-like prophet over any period of the last 500 years? With all the geographic and demographic facts of the time, were the rise of the Soviet Union in 1917 or collapse of France in 1940 foreseeable?

Or take Britain, which many post-Brexit are wishing to oblivion. You can look at its success in COVID vaccine development and rollout. Or you can look further back at its singular resolution against the Nazi conquest of Europe in World War II. Which of those would have been predicted in advance?

Czechoslovakia’s own version of the Maginot Line was a geographic fact you could have bet on against German invaders – yet Czechoslovakia was dismantled by political means. It’s geographic fortitude and threat to German invaders became irrelevant.

Political machinations defy geographic facts. So does technology defy would-be prophets. Technology changes the landscape, shuffles the strategic deck. These changes are inherently unforeseeable, or at least their timing. So the Dreadnaught relegated Britain’s naval supremacy to obsolescence ahead of WWI. Or motorised warfare changed what was possible in WWII. Or countless other advances in small and large ways irrevocably changed military and civilisational landscapes over time, allowing

one people to dominate another in new ways. This is another way in which Zeihan's geographic essentialism, casting the world into the pre-WWII world of regional powers, is deeply pessimistic. It implies a fundamental stasis in technology. Would another commentator in any other period have been correct even decades out, let alone a century? So what to make of Zeihan's brilliant claims today?

"This time is different" is a punchline in the finance industry. Because it's *never* different, is the implication. Except, sometimes, it is. Railways. Electricity. Combustible engines. The Internet. Maybe today we really have broken the bonds of geography. Maybe instant communication and the dawn of global, decentralised finance means new kinds of cities will emerge (or rather, be built, by new pioneers like [Balaji Srinivasan](#)). Maybe we still underestimate the Internet – much more has yet to move online. This vector in human history is new. It's anti-geography – it's everywhere. The closer we move towards this future, the further behind fall the inevitable. It's unsurprising that Balaji [disagrees](#) with Zeihan, preferring to bet on humanity's indomitable will to create its own future everywhere, and to leave the morass of US and global institutions in the past.

To the extent that Maçães is a prophet, he is less ambitious, his prophesy less potent. He can claim that he mainly outlines, in lovely detail, the blurring of lines on the world's greatest supercontinent. And who can really take issue with that?

Reversion vs Evolution

One way to frame Zeihan's thesis is that the post-WWII US World Order is an aberration in world history. As it recedes, so we will move towards a more normal state of nature – regional nation state competition. Zeihan's twenty first century is bleak, as the world reverts to a bleak historic mean. The world leading to WWII was unending chaos, a *realpolitik* morass of national states rising and falling in a race for self-preservation and dominance. And even the destruction of the world and coalescing of forces behind the US and USSR into the Cold War was in some ways a continuation of that: localised

proxy conflicts around the world that were really manifestations of two great powers doing battle.

And so, says Zeihan, this momentary blip in world history, where the seas are protected and global trade can flow, will recede. We will revert to a more normal world, of regional powers and local conflict, unfrozen from the stasis of global hegemony.

“It’ll be less like the messiness of the early 2000s or the raw potential of the 1950s, and more a disastrous combination of the battle royales and displacements of the 1870s against the economic backdrop of the 1930s.”

It is unsurprising then that Zeihan’s vision is such an analogue one. His twenty first century resembles the nineteenth, just with broadband and supercarriers. Game-changing technology gets a hand wave dismissal – artificial intelligence, or the kind that would matter, is nowhere. And that’s it. It all still comes down to geography. Infrastructure will still be prohibitive across African “*stacked plateaus*”. Brazil’s agriculture will still be sub-economic. Whatever liberal principles developed since WWII are for nothing or were window dressing to begin with.

Maçães’ future evolves rather than reverts. He proposes that “*we need a word to refer to the world as it is rather than the world as it aspires to be*” – and that word is Eurasia, and it is set to do “*what ‘Europe’ did for the Europeans after 1815.*” The world is evolving to its next phase, just as it arose from its prior state.

The Russian soul

If Maçães’ journey is part odyssey, his political wonderings recall Hannah Arendt. He whirlwinds through the histories of peoples and political philosophies, all tied with piercing insights, anecdotes and irony.

He prowls through the bowls of Russian historic originalism, overturning predecessors of Eurasianism, who tied Russia's past and future inextricably to Asia's, to the Mongols and the Turks. He paraphrases a Russian philosopher, referring to a "*subterranean affinity of souls between Russians and the Turkic peoples of the steppes*" – a very Russian conception of themselves. Arendt was dismissive of this self-conception:

"...to an innocent observer (as most Westerners were) the so-called Eastern soul appeared to be incomparably richer, its psychology more profound, its literature more meaningful than that of the "shallow" Western democracies...Franz Kafka knew well enough the superstition of fate which possesses people who live under the perpetual rule of accidents, the inevitable tendency to read a special superhuman meaning into happenings whose rational significance is beyond the knowledge and understanding of the concerned. He was well aware of the weird attractiveness of such peoples, their melancholy and beautifully sad folk tales which seemed so superior to the lighter and brighter literature of more fortunate peoples. He exposed the pride in necessity as such, even the necessity of evil, and the nauseating conceit which identifies evil and misfortune with destiny." (Origins of Totalitarianism)

Maçães ponders no less than the meaning of life and its relation to the state. He wraps novelists and poets and a nation's spirit into its geopolitical direction:

"[Russian poet Joseph] Brodsky thought it was to the credit of the dying Soviet government that it did not even try to evade, simplify or disguise the question [of how to live]. There was no answer, there is no meaning to life, and people had simply to live with that. As the novelist Victor Pelevin puts it in The Sacred Book of the Werewolf, the substance of human life actually changes very little from culture to culture, but human beings require a beautiful wrapper to cover it. Russian culture, uniquely, fails to provide one, and it calls this state of affairs 'spirituality'."

The language of Zeihan's prophesy can speak only to Russia's demographics and geographic sprawl sending it inexorably into chaos. Maçães contemplates the organs of the state and their relation with power:

"The state aspires not to overcome and replace chaos, but to nationalize it or, in other words, to acquire and enjoy monopoly of it."

Perhaps this added texture is just an indulgence. More likely, in such cultural excavations comparisons between the two do neither justice. Maçães is portending a new geopolitical future, but he is also doing a great many other things, including illuminating the very many puzzle pieces of disparate people with disparate histories. Prophecy is in some respects beside the point.

Lighting the path ahead

Neither book offers policy actions. Actions imply agency and do not fit Zeihan's deterministic approach. Maçães savours the world more as it is than what it will be – what are particular policies beside the grand arcs of peoples and their dreams? Zeihan's model is difficult to resist. Its weakness is precisely its strength: it cuts through the infinite strands of future possibilities and individual dreams to provide a holistic framework from which to assess a nation's strategic position. From there, policy makers can form insights.

I have been unable to resist applying it to my home Australia for example, largely untreated in *Disunited Nations*, presumably for its irrelevance to world affairs. Australia would do very well indeed against Zeihan's four criteria: ocean girt, internally stable and food and energy self-sufficient as it is, albeit with some qualms around its demographics (too small).

As a tour de force through the world's powers, their strategic positions and how they might unfold, *Disunited Nations* is exceptional. As prophesy, we shall see. But his warnings about darker days resonate uncomfortably. For me, as part of the generation raised through the golden 90's and '00's in the West, who believed history was behind them and its arc had bent for good to the justice of an iridescent, endless present, the ruptures in that fantasy align with the dusk of the US World Order. Whether or not it plays out in the way described, that our days and the days of our children may well be poorer and more fearful sits heavy.

Maçães might give reason to be more hopeful, for the future is unwritten. Thiel might say *of course* the future *may* be dark, for it will be whatever we make of it. Deutsch would say, well, everything that is permitted by the physical universe is possible, we just need the right knowledge. Maçães may in the end be the real conservative – it is far from radical to bet on the super-continent and its resident great powers that dominated world history until relatively recently. Whether at the heart of those lands or at their peripheries, eddy the stories and the people who will rise to the occasion or fall to be wondered over by future adventurers and storytellers.

Driven to Distraction: Recognizing and Coping with Attention Deficit Disorder from Childhood Through Adulthood by Edward M. Hallowell & John J. Ratey

After reading "Driven to Distraction",
I mostly try to avoid using the term "ADHD".
Hyperactivity is not uncommon in people with ADD
(Attention Deficit Disorder),
but it is not universal.
After reading the book,
I think lumping it in the definition of the syndrome
is doing a disservice to the people who have it --
especially those that do not exhibit hyperactivity.

Hypothetical Condition

Imagine that you got headaches every time things were too calm.
You would naturally try to do anything to avoid the calm.
As a kid, you would start misbehaving in class,
because kids have limited options.
As you grow up, you would find more interesting ways
to stave off the headaches.

Maybe you would take to driving 20 mph over the speed limit on the highway,
or take up skydiving as a hobby.
Maybe you could even turn your aversion to calmness into a
competitive advantage:
find a high-stress job like air traffic controller
or high-stakes financial analyst.
In college, you would worry about the due date for your paper,
but not do anything about it:
the worrying staves off the calm,
and finishing the paper in an all-nighter is super-stressful.

Attention Deficit Disorder is not like that;
it is much worse.

Instead of just getting a headache,
you are drowning in distractions without something to deliver
high levels of stimulation.

Fortunately,
the difficulty created by having such a challenging disorder
provides much needed stimulation.
People with ADD often use this stimulation to develop coping mechanisms.
Unfortunately, without a lot of luck or outside help,
those coping mechanisms are always stuck on a local maximum.
As soon as they work too well,
things calm down,
and they stop working.

One great way to feel better about yourself and the coping mechanisms
you have developed
is to read about other people,
and be grateful to whatever random forces steered you away
from coping mechanisms that are even more destructive.

Case Studies

The book has a lot of case studies.
I think the point is to show different facts of ADD,
or to show the diversity of the manifestations of the syndrome.
But, instead, I learned something else from them.
The horrifying truth about ADD is that coping mechanisms can always
get worse.

The case studies start mildly.
First, Jim, whose main coping mechanism is waiting until two hours
before week long projects are due to start working on them.
Luckily, his ideas are just good enough that even when he gets fired
from one job, he manages to find another one.
He doesn't get promoted, and often someone else gets the credit
for cleaning up his rough drafts, but at least
he survives.

The case studies continue with Carolyn,

who eventually became a therapist specializing in ADD, but as a kid, her main coping mechanism was running away from school, climbing on top of water towers and diving into a book.

The stimulus of climbing those rickety structures, plus the excitement of the plot-lines, helped her focus.

However, these coping skills were less helpful in avoiding the inevitable scolding of Carolyn's mother when, every Saturday, her mother learned of her antics.

Max, who earned the nickname Mad Max, used his nickname as a coping mechanism. He behaved in extreme ways, like intentionally tripping over his feet, and then pointing to himself and saying "crazy".

Danny and David, adopted twins, took to physical activity. Mostly in the form of fighting with each other, fighting with the teachers in the school for special children, and occasionally doing cartwheels when they were bored. What they got in return was a lot of physical restraint, to the point of harm, from the teachers.

The stories continue, not necessarily in order how destructive the coping behavior is which makes it all the more scary to read. You never know if you already read the worst one.

Jack, an editor, picks fights with everyone at work, and compensates for his egregious behavior by being very good at his job. George does all his work from his phone while driving, because that's the only way he can concentrate. Laura developed anxiety in order to focus herself. Douglas turned to alcoholism and workaholism.

Other People

No man is an island, and people with ADD are no different. Many people grow up in a family with parents and siblings.

Many people share their life with a long term partner.

The most emotional chapters in the book discuss the story of the people who live *near* ADD, especially undiagnosed ADD.

You cannot avoid having sympathy for the wife of someone who is qualified to be an emergency-room doctor, but then changed careers to a cartoonist, and then gave up on it to be a freelance writer.

But even harder to read was the story of Tommy and how his ADD organized his family's dynamics into complete dysfunction. Tommy's role, as is common for teenagers with ADD, was to be the screw-up. School is boring even for neurotypical teenagers, for people with ADD it looks designed as a way to make them fail.

Tommy's mother's role is to react with extreme emotions: compassion, worry, or anger, more or less in that order. When Tommy fails to respond to her compassion, she becomes worried, and when Tommy reacts, she responds with anger and pain.

Tommy's father's role is that of the harsh disciplinarian, giving harsh penalties when Tommy screws up or hurts his mother.

The role of Alex and Suzie, Tommy's siblings, is to serve as a contrast. Where he screws up, they toe the line. They listen to their parents. As their reward, they get to join their parents side in the big struggle the family is locked in.

When it is indicated that Tommy's problem might be treatable, everyone's roles is threatened. Not only does Tommy's ADD have to be addressed, but the entire family needs to learn how to function without the his ADD providing the stimulation for everyone else.

Treatment

As a book for lay people, those suffering from ADD or close to someone who does, the book does not focus on when to prescribe medicine.

It is pretty firm in saying that when medicine is prescribed,
it is a good idea to take it.
Further, the book stands up for informed consent,
and does stress the importance of
being informed:
often, the resistance to using medication is based on folk tales.

However,
while medication helps with focus,
it usually does not,
by itself,
help unlearn the harmful coping mechanisms.
Though unnecessary, those behaviors have often already become habit.
The book recommends therapy,
but also suggests that other lay people can help
the person with ADD figure out,
and keep to,
structures that will help them.

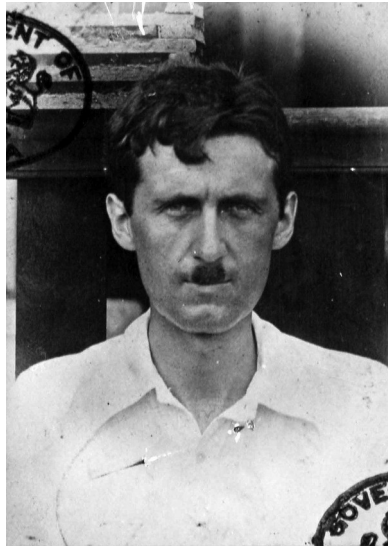
Conclusion

As with any psychological condition,
and as seems to be the standard in these books,
a list of great people who also suffered from the condition
is mentioned.
Like Robin Williams as the mascot of high-functioning depression,
people with ADD get Albert Einstein.

It is nice to see that the condition does
not have to be debilitating.
Some people with the condition have made
incredible accomplishments.
It is also important to remember that there are a lot of people
with ADD who are not Einstein,
and should not feel bad about it.

Down and Out in Paris and London by George Orwell

George Orwell's *Down and Out in Paris and London* is at least three things; a highly entertaining, almost picaresque tale of rough-and-tumble living in Europe, a serious attempt to catalogue the numerous humiliations and injustices impoverished people were exposed to in Orwell's time, and a stark comparison between life as a tramp who makes use of robust, if hellish and kafkaesque welfare resources, and as one who tries to get by working terrible jobs and living in disgusting places.



Above we see a young(ish) Orwell sporting a very Orwellian mustache.

Orwell begins in Paris, where a long period of intermittent employment followed by a robbery have reduced him from poor-to-middle-class-expat to actual poor person. This is Orwell's first published book, and he's quite young (23-25) at the time of writing. I mention this only because, suprisingly, Orwell utterly lacks romantic notions about living an impoverished, bohemian life in a world-class city. Instead he characterizes his plunge into the Parisian underworld as a means of purging himself of the prejudices he aquired as an upper middle class Etonian. His descriptions of the characters in the run down hotel where he starts out are about as close as he gets to Kerouacish gushing about the wacky beatitude that arises out of a life in poverty:

"There were eccentric characters in the hotel. The Paris slums are a gathering-place for eccentric people—people who have fallen into solitary, half-mad grooves of life and given up trying to be normal or decent. Poverty frees them from ordinary standards of

behaviour, just as money frees people from work. Some of the lodgers in our hotel lived lives that were curious beyond words.

There were the Rougiers, for instance, an old, ragged, dwarfish couple who plied an extraordinary trade. They used to sell postcards on the Boulevard St Michel. The curious thing was that the postcards were sold in sealed packets as pornographic ones, but were actually photographs of chateaux on the Loire; the buyers did not discover this till too late, and of course never complained. The Rougiers earned about a hundred francs a week, and by strict economy managed to be always half starved and half drunk. The filth of their room was such that one could smell it on the floor below. According to Madame F., neither of the Rougiers had taken off their clothes for four years."

I can't resist including another of these little characterizations. It's amazing what Orwell can pack into a paragraph:

"Or there was Henri, who worked in the sewers. He was a tall, melancholy man with curly hair, rather romantic-looking in his long, sewer-man's boots. Henri's peculiarity was that he did not speak, except for the purposes of work, literally for days together. Only a year before he had been a chauffeur in good employ and saving money. One day he fell in love, and when the girl refused him he lost his temper and kicked her. On being kicked the girl fell desperately in love with Henri, and for a fortnight they lived together and spent a thousand francs of Henri's money. Then the girl was unfaithful; Henri planted a knife in her upper arm and was sent to prison for six months. As soon as she had been stabbed the girl fell more in love with Henri than ever, and the two made up their quarrel and agreed that when Henri came out of jail he should buy a taxi and they would marry and settle down. But a fortnight later the girl was unfaithful again, and when Henri came out she was with child, Henri did not stab her again. He drew out all his savings and went on a drinking-bout that ended in another month's imprisonment; after that he went to work in the sewers. Nothing would induce Henri to talk. If you asked him why he worked in the sewers he never answered, but simply crossed his wrists to signify handcuffs, and jerked his head southward, towards the prison. Bad luck seemed to have turned him half-witted in a single day."

Orwell rarely stops to inject political or sociological speculation. Rather he tells the story as it happened, and keeps his more academic conclusions to just a couple chapters at the end. I'm going to attempt to honor this structure by going through the book chronologically, and then analyzing his actual arguments for and against such-and-such-solution at the very end.

Orwell's greatest strength as a writer, wry detachment, also tends to be greatest weakness. Occasionally it feels like his disgust, or else his commitment to a somewhat journalistic tone stop him from confronting the sheer grotesquity of some of the people and situations he's met with. Take his profile of fellow British ex-pat Charlie. He recalls Charlie getting drunk and raping a young prostitute, and then later rhapsodizing over the experience:

“And so, just for one instant, I captured the supreme happiness, the highest and most refined emotion to which human beings can attain. And in the same moment it was finished, and I was left—to what? All my savagery, my passion, were scattered like the petals of a rose. I was left cold and languid, full of vain regrets; in my revulsion I even felt a kind of pity for the weeping girl on the floor. Is it not nauseous, that we should be the prey of such mean emotions? I did not look at the girl again; my sole thought was to get away. I hastened up the steps of the vault and out into the street. It was dark and bitterly cold, the streets were empty, the stones echoed under my heels with a hollow, lonely ring. All my money was gone, I had not even the price of a taxi fare. I walked back alone to my cold, solitary room...but there, messieurs et dames, that is what I promised to expound to you. That is Love. That was the happiest day of my life”

And what does George ‘Conscience of His Generation’ Orwell have to say about this?:

“He was a curious specimen, Charlie. I describe him, just to show what diverse characters could be found flourishing in the Coq d’Or quarter.”

Dissappointing, to say the least. I wonder if Orwell is attempting to be funny, using his callousness to reflect the callousness of everyone who sits around sipping their drinks as Charlie tells poetic stories about raping prostitutes. Or perhaps he thought the story spoke for itself and required no further comment. Luckily Orwell doesn’t do this very often. Most of the time he holds the camera on scenes of degradation and injustice far past the point where we as the reader would prefer to look away, and the book is all the better for it. Such scenes aren’t necessarily disgusting or dirty. Instead, Orwell indulges in an almost Nietzschean interest in the psychological impact of petty humiliations:

“You discover, for instance, the secrecy attaching to poverty. At a sudden stroke you have been reduced to an income of six francs a day. But of course you dare not admit it—you have got to pretend that you are living quite as usual. From the start it tangles you in a net of lies, and even with the lies you can hardly manage it. You stop sending clothes to the laundry, and the laundress catches you in the street and asks you why; you mumble something, and she, thinking you are sending the clothes elsewhere, is your enemy for life. The tobacconist keeps asking why you have cut down your smoking. There are letters you want to answer, and cannot, because stamps are too expensive. And then there are your meals—meals are the worst difficulty of all. Every day at meal-times you go out, ostensibly to a restaurant, and loaf an hour in the Luxembourg Gardens, watching the pigeons. Afterwards you smuggle your food home in your pockets. Your food is bread and margarine, or bread and wine, and even the nature of the food is governed by lies. You have to buy rye bread instead of household bread, because the rye loaves, though dearer, are round and can be smuggled in your pockets. This wastes you a franc a day. Sometimes, to keep up appearances, you have to spend sixty centimes on a drink, and go correspondingly short of food. Your linen gets filthy, and you run out of soap and razor-blades. Your hair wants cutting, and you try to cut it yourself, with such fearful results that you have to go to the barber after all, and spend the equivalent of a day’s food. All day you are telling lies, and expensive

lies...one could multiply these disasters by the hundred. They are part of the process of being hard up."

You might say these humiliations stem from Orwell's sudden fall from low-middle-class respectability. But later Orwell will stress that proud, life-long tramps are fairly rare. Most of the people he encounters were also respectable low-to-middle-class people once. So this fall from grace and the obsession with some degree of keeping up appearances is shared by most tramps. Later, when Orwell watches his friend Boris prepare for a job hunt, we see the level of skill some tramps(I'm going to use this word as Orwell does) have acquired in transforming themselves into phantoms of their respectable pasts:

"All the clothes he now had left were one suit, with one shirt, collar and tie, a pair of shoes almost worn out, and a pair of socks, all holes. He had also an overcoat which was to be pawned in the last extremity. He had a suitcase, a wretched twenty-franc cardboard thing, but very important, because the patron of the hotel believed that it was full of clothes—without that, he would probably have turned Boris out of doors. What it actually contained were the medals and photographs, various odds and ends, and huge bundles of love-letters. In spite of all this Boris managed to keep a fairly smart appearance. He shaved without soap and with a razor-blade two months old, tied his tie so that the holes did not show, and carefully stuffed the soles of his shoes with newspaper. Finally, when he was dressed, he produced an ink-bottle and inked the skin of his ankles where it showed through his socks. You would never have thought, when it was finished, that he had recently been sleeping under the Seine bridges."

Orwell and Boris share a pretty cute bromance for most of the Parisian section, spending nights together in terrible cheap rooms, discussing their future prospects. Usually, Boris(a Russian expat whose parents were murdered by the Bolsheviks) serves as the brains of the operation:

"What things a man can do with brains! Brains will make money out of anything. I had a friend once, a Pole, a real man of genius; and what do you think he used to do? He would buy a gold ring and pawn it for fifteen francs. Then—you know how carelessly the clerks fill up the tickets—where the clerk had written "en or" he would add "et diamants" and he would change "fifteen francs" to "fifteen thousand". Neat, eh? Then, you see, he could borrow a thousand francs on the security of the ticket. That is what I mean by brains...For the rest of the evening Boris was in a hopeful mood, talking of the times we should have together when we were waiters together at Nice or Biarritz, with smart rooms and enough money to set up mistresses. He was too tired to walk the three kilometres back to his hotel, and slept the night on the floor of my room, with his coat rolled round his shoes for a pillow."

The two of them carry on like this for some time, drifting aimlessly though Paris, hoping for work. At one point they get scammed by a fake Bolshevik cell that disappears once they pay their membership dues. Orwell's underrated and understated sense of humor

is on full display after he and Boris return to the cell office to find it deserted and totally sans Lenin posters:

“And that was the last we ever heard of the secret society. Who or what they really were, nobody knew. Personally I do not think they had anything to do with the Communist Party; I think they were simply swindlers, who preyed upon Russian refugees by extracting entrance fees to an imaginary society. It was quite safe, and no doubt they are still doing it in some other city. They were clever fellows, and played their part admirably. Their office looked exactly as a secret Communist office should look, and as for that touch about bringing a parcel of washing, it was genius.”

Eventually Orwell does manage to find work, as a scullion in a large, upscale hotel's underground kitchens, and so begins the *Kitchen Confidential* section of the book. I love that Orwell feels free to devote such a big section of the story to describing his day-to-day in this horrible, hellish job. There seems to be no doubt in his mind that readers would find it interesting. And his writerly instincts are soon proven out. I first read this book years ago, and whenever it randomly comes to mind, it's because of these images of Orwell the scullion. Here he is going down into the depths of the hotel for the first time:

“He led me down a winding staircase into a narrow passage, deep underground, and so low that I had to stoop in places. It was stiflingly hot and very dark, with only dim, yellow bulbs several yards apart. There seemed to be miles of dark labyrinthine passages—actually, I suppose, a few hundred yards in all—that reminded one queerly of the lower decks of a liner; there were the same heat and cramped space and warm reek of food, and a humming, whirring noise (it came from the kitchen furnaces) just like the whir of engines. We passed doorways which let out sometimes a shouting of oaths, sometimes the red glare of a fire, once a shuddering draught from an ice chamber. As we went along, something struck me violently in the back. It was a hundred-pound block of ice, carried by a blue-aproned porter. After him came a boy with a great slab of veal on his shoulder, his cheek pressed into the damp, spongy flesh. They shoved me aside with a cry of 'Sauve-toi, idiot!' and rushed on. On the wall, under one of the lights, someone had written in a very neat hand: 'Sooner will you find a cloudless sky in winter, than a woman at the Hôtel X who has her maidenhead.' It seemed a queer sort of place.”

The 'it seemed a queer sort of place' succeeds at Monty Pythonesque humor where his similar treatment of Charlie the rapist seemed to fail. Or perhaps Orwell isn't being funny, and I think of this attitude as Monty Pythonesque because Python sought to mock the stiff upper lip types of Orwell's generation. Who knows? I don't know enough of the nuances of British humor to be sure.

Orwell's work as a scullion turns out to be less a merciful reprieve from the life of a starving tramp, and more a daily tour of hell. He works 7 am to 9 pm everyday except Sunday, with a work flow like this:

"I calculated that one had to walk and run about fifteen miles during the day, and yet the strain of the work was more mental than physical. Nothing could be easier, on the face of it, than this stupid scullion work, but it is astonishingly hard when one is in a hurry. One has to leap to and fro between a multitude of jobs—it is like sorting a pack of cards against the clock. You are, for example, making toast, when bang! down comes a service lift with an order for tea, rolls and three different kinds of jam, and simultaneously bang! down comes another demanding scrambled eggs, coffee and grapefruit; you run to the kitchen for the eggs and to the dining-room for the fruit, going like lightning so as to be back before your toast bums, and having to remember about the tea and coffee, besides half a dozen other orders that are still pending; and at the same time some waiter is following you and making trouble about a lost bottle of soda-water, and you are arguing with him. It needs more brains than one might think. Mario said, no doubt truly, that it took a year to make a reliable cafetier."

On top of that, there's a quaint(in retrospect)race hatred between all the employees, who tend to be sorted into their various positions by ethnicity:

"The office employees and the cooks and sewing-women were French, the waiters Italians and Germans (there is hardly such a thing as a French waiter in Paris), the plongeurs of every race in Europe, beside Arabs and Negroes. French was the lingua franca, even the Italians speaking it to one another."

Luckily, none of it seems to matter much once orders stop coming in and everyone is indiscriminately screaming insults at everyone else. If you've ever seen on of those Gordon Ramsay cooking shows, what Orwell describes is no different. He admires the order that emerges from the chaos of dozens of people who share no common background or language screaming at one another to hurry up with the damn sauce already. And in his words, all the hustle and hubbub is 'the good side of hotel work'.

So what's the bad side?

"-it is this—that the job the staff are doing is not necessarily what the customer pays for. The customer pays, as he sees it, for good service; the employee is paid, as he sees it, for the boulot—meaning, as a rule, an imitation of good service. The result is that, though hotels are miracles of punctuality, they are worse than the worst private houses in the things that matter.

Take cleanliness, for example. The dirt in the Hôtel X, as soon as one penetrated into the service quarters, was revolting. Our cafeteria had year-old filth in all the dark corners, and the bread-bin was infested with cockroaches. Once I suggested killing these beasts to Mario. 'Why kill the poor animals?' he said reproachfully. The others laughed when I wanted to wash my hands before touching the butter. Yet we were clean where we recognized cleanliness as part of the boulot. We scrubbed the tables and polished the brasswork regularly, because we had orders to do that; but we had no orders to be genuinely clean, and in any case we had no time for it. We were simply

carrying out our duties; and as our first duty was punctuality, we saved time by being dirty."

I think we SSC readers might jump at the chance to call this 'a misalignment between the real interests of the customers and the incentives of the employees, or else as an 'optimizing for optics over that which those optics are presumed to represent' (though maybe that second one is redundant). But neither of those reframings cast much light on the situation for me. Instead I'm left wondering at the real value of kitchen cleanliness as such, in contrast to 'cleanliness' as measured by number of people made sick by food from that kitchen. My thinking is that all the grime in the world shouldn't matter so long as it doesn't result in any form of customer dissatisfaction, but then so much of that dissatisfaction is determined by how much information a customer gets about the process of food preparation. It's a tree falling in the forest problem, isn't it? Cleanliness matters if you have means to judge cleanliness. If you have no means, the way your food is made hardly matters at all (I guess that makes me a consequentialist?) This leads me to wonder if the popularity of television programs showing how processed foods are produced have any effect on their performance in the market.

But then, that's the funny thing about cleanliness: a person who eats big macs probably has a vague understanding of the way that they're made, and that it isn't a very attractive process. They know that frozen patties are placed in a heating tray. That's all well and good. But god forbid there's a bit of *grime* on the tray that heats up the processed meat paddy...now that would be unforgivable. I get the sense that people who lived before the mid-twentieth century had opposite feelings about cleanliness: if your chicken shank falls in the mud, who cares? Wipe it off and eat it. But if that chicken led a wretched life and fell sick before it was slaughtered, you'd better be careful. We moderns don't seem to give a damn about that kind of thing, so long as our notions of visible cleanliness are maintained. We make a good foil for Orwell's hotel patrons. But I digress:

"Dirtiness is inherent in hotels and restaurants, because sound food is sacrificed to punctuality and smartness. The hotel employee is too busy getting food ready to remember that it is meant to be eaten. A meal is simply 'une commande' to him, just as a man dying of cancer is simply 'a case' to the doctor. A customer orders, for example, a piece of toast. Somebody, pressed with work in a cellar deep underground, has to prepare it. How can he stop and say to himself, 'This toast is to be eaten—I must make it eatable'? All he knows is that it must look right and must be ready in three minutes. Some large drops of sweat fall from his forehead on to the toast. Why should he worry? Presently the toast falls among the filthy sawdust on the floor. Why trouble to make a new piece? It is much quicker to wipe the sawdust off. On the way upstairs the toast falls again, butter side down. Another wipe is all it needs. And so with everything. The only food at the Hôtel X which was ever prepared cleanly was the staff's, and the patron's. The maxim, repeated by everyone, was: 'Look out for the patron, and as for the clients, s'en f—pas mal!' Everywhere in the service quarters dirt festered—a

secret vein of dirt, running through the great garish hotel like the intestines through a man's body."

Earlier I mentioned that this section is reminiscent of Anthony Bourdain's *Kitchen Confidential*, a book whose very success proves that we cleanliness obsessed moderns get a voyeuristic kick out of stories such as Orwell's. I think it's this same voyeuristic kick that makes *Down and Out* an appealing book in general. Just as we all go to restaurants and know almost nothing of what goes on behind kitchen doors, all of us see tramps most everyday and know little about how they live and think. And just like in the case of the kitchens, as much as we get a kick of being let in on the secrets of the unknown underworld, we also have very little desire to actively seek out that information for ourselves. Because a lot of the time, that information hurts.

Orwell ends this rather disturbing expose of hotel restaurants with some good-natured and utterly English needling of the American palate:

"According to Boris, the same kind of thing went on in all Paris hotels, or at least in all the big, expensive ones. But I imagine that the customers at the Hôtel X were especially easy to swindle, for they were mostly Americans, with a sprinkling of English—no French—and seemed to know nothing whatever about good food. They would stuff themselves with disgusting American 'cereals', and eat marmalade at tea, and drink vermouth after dinner, and order a poulet à la reine at a hundred francs and then souse it in Worcester sauce. One customer, from Pittsburgh, dined every night in his bedroom on grape-nuts, scrambled eggs and cocoa. Perhaps it hardly matters whether such people are swindled or not."

This, at least, was comforting. As were all the colorful tales of criminal hijinks:

"There were tales of dope fiends, of old debauchees who frequented hotels in search of pretty page boys, of thefts and blackmail. Mario told me of a hotel in which he had been, where a chambermaid stole a priceless diamond ring from an American lady. For days the staff were searched as they left work, and two detectives searched the hotel from top to bottom, but the ring was never found. The chambermaid had a lover in the bakery, and he had baked the ring into a roll, where it lay unsuspected until the search was over."

And despite all of that, Orwell says he was reasonably satisfied with his life at the time. Anyone who has lived the life of a manual laborer, even for a very short time, will be familiar with this feeling:

"I had no sensation of poverty, for even after paying my rent and setting aside enough for tobacco and journeys and my food on Sundays, I still had four francs a day for drinks, and four francs was wealth. There was—it is hard to express it—a sort of heavy contentment, the contentment a well-fed beast might feel, in a life which had become so simple. For nothing could be simpler than the life of a plongeur. He lives in a rhythm

between work and sleep, without time to think, hardly conscious of the exterior world; his Paris has shrunk to the hotel, the Métro, a few bistros and his bed. If he goes afield, it is only a few streets away, on a trip with some servant-girl who sits on his knee swallowing oysters and beer. On his free day he lies in bed till noon, puts on a clean shirt, throws dice for drinks, and after lunch goes back to bed again. Nothing is quite real to him but the boulot, drinks and sleep; and of these sleep is the most important."

Yet I don't quite believe that Orwell, a highly educated man, and, well, *George Orwell*, was quite as content as he makes himself out to be. If there is any privilege or classist blindspot that Orwell himself fails to acknowledge here, it is that throughout this book, there is a sense of inevitability around Orwell's eventual escape from this life. I can't point at a specific passage where this makes itself apparent, but I find it permeates the entire text; Orwell knows that eventually he will get out of this, someday, somehow. The bloke went to *Eton*, after all. That will come up again later on, but suffice it to say, Orwell's situation never feels really, truly desperate to me. But perhaps that's just testament to his complete and utter lack of self-pity...and to the intense but fleeting pleasures of a working class life, characterized by back breaking work punctuated by long bouts of drinking:

"By half past one the last drop of pleasure had evaporated, leaving nothing but headaches. We perceived that we were not splendid inhabitants of a splendid world, but a crew of underpaid workmen grown squalidly and dismally drunk. We went on swallowing the wine, but it was only from habit, and the stuff seemed suddenly nauseating. One's head had swollen up like a balloon, the floor rocked, one's tongue and lips were stained purple. At last it was no use keeping it up any longer. Several men went out into the yard behind the bistro and were sick. We crawled up to bed, tumbled down half dressed, and stayed there ten hours.

Most of my Saturday nights went in this way. On the whole, the two hours when one was perfectly and wildly happy seemed worth the subsequent headache. For many men in the quarter, unmarried and with no future to think of, the weekly drinking-bout was the one thing that made life worth living."

Never underestimate the power a regular and reliable debauch can have on the endurance of people without prospects. For most of human history, it seems to have been enough to keep peasants pushing their ploughs. At least, that's what I get from Orwell's descriptions of the roughly six-hours-a-week of bacchanalian fun he and the other residents of his quarter seemed to have attended with almost religious fervor and regularity.

Time wears on, and Orwell finally gets a position at a newly opened restaurant that he'd been promised several months previous. The most notable character here is the ardent-communist-by-affiliation-and-temperment whom Orwell(a socialist himself) regards with passing interest. Here's him scolding Orwell for daring to work:

“Put that brush down, you fool! You and I belong to proud races; we don't work for nothing, like these damned Russian serfs. I tell you, to be cheated like this is torture to me. There have been times in my life, when someone has cheated me even of five sous, when I have vomited—yes, vomited with rage.

'Besides, mon vieux, don't forget that I'm a Communist. À bas la bourgeoisie! Did any man alive ever see me working when I could avoid it? No. And not only that I don't wear myself out working, like you other fools, but I steal, just to show my independence. Once I was in a restaurant where the patron thought he could treat me like a dog. Well, in revenge I found out a way to steal milk from the milk-cans and seal them up again so that no one should know. I tell you I just swilled that milk down night and morning. Every day I drank four litres of milk, besides half a litre of cream. The patron was at his wits' end to know where the milk was going. It wasn't that I wanted milk, you understand, because I hate the stuff; it was principle, just principle.’

This new job, still as a scullion, but now in a small restaurant frequented by Russian expats(who, because this is 1920s Paris, have all lived extraordinary, tragic, lives) turns out to be even worse than the one at the hotel. He works from:

“...seven in the morning till half past twelve the next morning—seventeen and a half hours, almost without a break. We never had time to sit down till five in the afternoon, and even then there was no seat except the top of the dustbin. Boris, who lived near by and had not to catch the last Métro home, worked from eight in the morning till two the next morning—eighteen hours a day, seven days a week. Such hours, though not usual, are nothing extraordinary in Paris.”

I found this almost too extreme to believe, simply because of the limits of the human body...until I moved to Japan. Now it seems a given that so long as 14+ hour working days are fairly ubiquitous, they will be endured without hesitation or complaint. On the surface, the causes of overwork in 1920s Paris and 21st century Japan appear quite different, but in the end they're basically the same: people work this way because if they don't, they'll soon be replaced by someone who will. Only in the last couple of decades has Japan realized that more than a labor issue, this is a public health crisis: people die if they carry on working like this for too long, even very young people(and the proof is in the pudding—most of my male high school students here in Japan have noticeable greying in their hair. I'm talking 15-17 year olds.)

But of course in Orwell's Paris, no one kept track of that sort of thing. Yet such a life does have it's compensations:

“At half past twelve I would put on my coat and hurry out. The patron, bland as ever, would stop me as I went down the alley-way past the bar. 'Mais, mon cher monsieur, how tired you look! Please do me the favour of accepting this glass of brandy.’

He would hand me the glass of brandy as courteously as though I had been a Russian duke instead of a plongeur. He treated all of us like this. It was our compensation for working seventeen hours a day."

I suppose this is the plaque-and-a-rolax of the restaurant life. Unfortunately, it's not quite enough to keep Orwell working seventeen hour days in a cramped, dirty kitchen. Soon he writes to a friend in England and begs for a job, and actually receives a reply that sends Orwell into daydreams about a more leisurely sort of employment:

"[I] was to look after a congenital imbecile, which sounded a splendid rest cure after the Auberge de Jehan Cottard. I pictured myself loafing in the country lanes, knocking thistle-heads off with my stick, feeding on roast lamb and treacle tart, and sleeping ten hours a night in sheets smelling of lavender."

But before moving on to England, Orwell gives a chapter of analysis of his life as a scullion, and what its greater social significance might be. He quickly labels the Paris scullions a class of modern slaves, and then wonders why such horrible conditions are allowed to exist in a city such as Paris. His answer is interesting: he sees the problem as basically one of public complacency. The people of Europe see the scullions' work as something that simply must be done, much like work in the sewers, coal mines etc. But Orwell doesn't see the scullion's work that way, at least, not in its current state:

"...it does not follow that he is doing anything useful; he may be only supplying a luxury which, very often, is not a luxury."

As an example of what I mean by luxuries which are not luxuries, take an extreme case, such as one hardly sees in Europe. Take an Indian rickshaw puller, or a gharry pony. In any Far Eastern town there are rickshaw pullers by the hundred, black wretches weighing eight stone, clad in loin-cloths. Some of them are diseased; some of them are fifty years old. For miles on end they trot in the sun or rain, head down, dragging at the shafts, with the sweat dripping from their grey moustaches. When they go too slowly the passenger calls them bahinchut. They earn thirty or forty rupees a month, and cough their lungs out after a few years. The gharry ponies are gaunt, vicious things that have been sold cheap as having a few years' work left in them. Their master looks on the whip as a substitute for food. Their work expresses itself in a sort of equation—whip plus food equals energy; generally it is about sixty per cent whip and forty per cent food. Sometimes their necks are encircled by one vast sore, so that they drag all day on raw flesh. It is still possible to make them work, however; it is just a question of thrashing them so hard that the pain behind outweighs the pain in front. After a few years even the whip loses its virtue, and the pony goes to the knacker. These are instances of unnecessary work, for there is no real need for gharries and rickshaws; they only exist because Orientals consider it vulgar to walk. They are luxuries, and, as anyone who has ridden in them knows, very poor luxuries. They afford a small amount of convenience, which cannot possibly balance the suffering of the men and animals."

Orwell served for five years as an imperial policeman in Burma(now Myanmar), and so isn't speaking from some vague antipathy towards Asian customs here. He saw such sights everyday for years on end, and it's no surprise that he would question their essential usefulness. But as brutal as the practice sounds, I find myself questioning his conclusion. Parts of India are punishingly hot, and many of the cities are large and spread out. It makes perfect sense to me that in a time before the wide spread use of cars, another cheap means of transportation would arise to meet the needs of the upper and middle class(because the rides are so cheap, they are not just a privilege of rulers)who want to avoid trudging miles across a large city in the middle of summer. To imply that rickshaws should be banned or phased out because they afford only "a small amount of convenience" strikes me as overreaching at best, and outright authoritarian at worst. But I find Orwell's basic argument much more compelling when applied to his own situation:

"Similarly with the plongeur. He is a king compared with a rickshaw puller or a gharry pony, but his case is analogous. He is the slave of a hotel or a restaurant, and his slavery is more or less useless. For, after all, where is the real need of big hotels and smart restaurants? They are supposed to provide luxury, but in reality they provide only a cheap, shoddy imitation of it. Nearly everyone hates hotels. Some restaurants are better than others, but it is impossible to get as good a meal in a restaurant as one can get, for the same expense, in a private house. No doubt hotels and restaurants must exist, but there is no need that they should enslave hundreds of people. What makes the work in them is not the essentials; it is the shams that are supposed to represent luxury. Smartness, as it is called, means, in effect, merely that the staff work more and the customers pay more; no one benefits except the proprietor, who will presently buy himself a striped villa at Deauville. Essentially, a 'smart' hotel is a place where a hundred people toil like devils in order that two hundred may pay through the nose for things they do not really want. If the nonsense were cut out of hotels and restaurants, and the work done with simple efficiency, plongeurs might work six or eight hours a day instead often or fifteen."

I see Orwell's dream as having come true in a limited sense. Upscale restaurants that serve terrible food made by people working for pennies are relatively rare now(except on cruise ships). Their two primary aspects (hellish working conditions and bad food)have split and diverged and now exist in different food industry niches: on the one hand we have the modern fast food restaurant, where the work is "done with simple efficiency...[scullions] might work six or eight hours a day" and true mid-to-upscale restaurants, where a higher level of cleanliness and quality in preparation and ingredients is assumed, and I think in most cases, delivered upon. In these sorts of places, which Bourdain describes in Kitchen Confidential, the chefs work in hellish conditions, but are paid relatively well and work more reasonable, if irregular and nocturnal, hours. In fast food places, people are paid less but are subject to a form of basic protection borne out of the sheer size and visibility of corporations like Burger King and McDonalds. No doubt Orwell would find this situation ghastly in it's own way, but I doubt he'd deny the life of restaurant and hotel workers has markedly improved since the 1920s.

Trouble is, Orwell doesn't see these horrible working conditions are merely a result of people's misguided desire to eat overpriced, low-quality food:

"I am trying to go beyond the immediate economic cause, and to consider what pleasure it can give anyone to think of men swabbing dishes for life. For there is no doubt that people—comfortably situated people—do find a pleasure in such thoughts...I believe that this instinct to perpetuate useless work is, at bottom, simply fear of the mob. The mob (the thought runs) are such low animals that they would be dangerous if they had leisure; it is safer to keep them too busy to think. A rich man who happens to be intellectually honest, if he is questioned about the improvement of working conditions, usually says something like this:

'We know that poverty is unpleasant; in fact, since it is so remote, we rather enjoy harrowing ourselves with the thought of its unpleasantness. But don't expect us to do anything about it. We are sorry for you lower classes, just as we are sorry for a cat with the mange, but we will fight like devils against any improvement of your condition. We feel that you are much safer as you are. The present state of affairs suits us, and we are not going to take the risk of setting you free, even by an extra hour a day. So, dear brothers, since evidently you must sweat to pay for our trips to Italy, sweat and be damned to you.'

And once again, my experiences in Japan stop me from dismissing Orwell's rather sinister notion of upper-classes motives outright. The following situation has occurred too many times:

Me: So why do Japanese people work so much? Teachers in the US usually go home at 4 or 5. Why do you stay until ten or eleven?

Japanese Colleague: Well, you know, I have to stay because everyone else does. It would look bad if I was first to leave.

Me: Okay, but what if that social pressure was removed?

Colleague: *Expression of shock and horror* but then, I wouldn't know what to do. I think people need to work, you know. If people have too much time, well...

Please don't think I'm exaggerating about Japan. I'm really, really, not. But in contrast to Orwell, I'd assert that everyone, including the mob themselves share this same fear of the mob. Everyone is afraid of too much improvement, if that improvement means more 'idle' time for poor people. We can spot this now whenever UBI is brought up. I count myself among the skeptics of UBI, but I recognize that much of my initial skepticism amounted to, "But what will all those people do without their horrible jobs to keep them busy?" And I see this mob-fear in the eyes of my Japanese friends whenever

I question them. They would probably still see it in my eyes if they asked me about UBI. Orwell sums up the whole, messy, tangled situation far better than I could:

“A plongeur is a slave, and a wasted slave, doing stupid and largely unnecessary work. He is kept at work, ultimately, because of a vague feeling that he would be dangerous if he had leisure. And educated people, who should be on his side, acquiesce in the process, because they know nothing about him and consequently are afraid of him. I say this of the plongeur because it is his case I have been considering; it would apply equally to numberless other types of worker. These are only my own ideas about the basic facts of a plongeur's life, made without reference to immediate economic questions, and no doubt largely platitudes. I present them as a sample of the thoughts that are put into one's head by working in an hotel.”

And with that, on to London!

II

After a long, mostly uneventful third-class trip by boat, Orwell arrives at his friend's office, bright-eyed and bushy-tailed about his future prospects as an imbecile chaperone:

“...and his first words knocked everything to ruins. 'I'm sorry,' he said; 'your employers have gone abroad, patient and all. However, they'll be back in a month. I suppose you can hang on till then?'

I was outside in the street before it even occurred to me to borrow some more money. There was a month to wait, and I had exactly nineteen and sixpence in hand. The news had taken my breath away. For a long time I could not make up my mind what to do. I loafed the day in the streets, and at night, not having the slightest notion of how to get a cheap bed in London, I went to a 'family' hotel, where the charge was seven and sixpence. After paying the bill I had ten and twopence in hand.”

Using [this converter](#) I see that ten and two pence in 1930 is equivalent to roughly 32 (2017) US dollars. Our friend Orwell was in a pretty tight spot. Yet it was never quite so tight as it seems. From my edition's introduction:

“When Orwell left Paris in December 1929 he did not, in fact, immediately live as a down-and-out in London. Instead, he spent Christmas with his family, whose joy was confined when their penniless son—now aged twenty-six and seemingly an unqualified failure—suddenly reappeared. Defensively he announced to all and sundry that he was working on a book about his time in Paris. But meanwhile he had somehow to earn something and tutoring jobs were found for him near Southwold. Also, he soon began to establish a reputation as a courageously independent-minded reviewer who was not overawed by such 'Big Names' as Edith Sitwell or J.B. Priestley.”

From this and other research into Orwell's life at this time, I think we can surmise that his life in Paris was in no way a performance, a LARP, or even an intentional bit of journalism, at least at first. Orwell simply had no other options. On the other hand, his life in London was, at least somewhat, chosen. Orwell's first published work would end up being an essay on conditions in 'Spikes' or, what basically amounted to state run shelters for tramps. But he had family willing to take him in, and shameful though it would be for a twenty five year old Eton graduate to resort to that, there was no real need for him to live as he did in London. But he did live in that way, and I think his reasons are mostly immaterial in regards to his experiences and insights. When class comes into play, Orwell readily acknowledges it. Yet it turns out to be less a factor than he'd first expected. He rarely receives special treatment, but it only adds to his sense of personal degradation:

"...I dared not speak to anyone, imagining that they must notice a disparity between my accent and my clothes. (Later I discovered that this never happened.) My new clothes had put me instantly into a new world. Everyone's demeanour seemed to have changed abruptly. I helped a hawkker pick up a barrow that he had upset. 'Thanks, mate,' he said with a grin. No one had called me mate before in my life—it was the clothes that had done it. For the first time I noticed, too, how the attitude of women varies with a man's clothes. When a badly dressed man passes them they shudder away from him with a quite frank movement of disgust, as though he were a dead cat. Clothes are powerful things. Dressed in a tramp's clothes it is very difficult, at any rate for the first day, not to feel that you are genuinely degraded. You might feel the same shame, irrational but very real, your first night in prison."

He finds cheap lodging in a "kip" wherein eight men sleep in a room fifteen feet square by eight high. As vivid and horrible as his descriptions of the subterranean kitchen, I somehow find this even worse:

"When I got into the bed I found that it was as hard as a board, and as for the pillow, it was a mere hard cylinder like a block of wood. It was rather worse than sleeping on a table, because the bed was not six feet long, and very narrow, and the mattress was convex, so that one had to hold on to avoid falling out. The sheets stank so horribly of sweat that I could not bear them near my nose. Also, the bedclothes only consisted of the sheets and a cotton counterpane, so that though stuffy it was none too warm. Several noises recurred throughout the night. About once in an hour the man on my left—a sailor, I think—woke up, swore vilely, and lighted a cigarette. Another man, victim of a bladder disease, got up and noisily used his chamber-pot half a dozen times during the night. The man in the corner had a coughing fit once in every twenty minutes, so regularly that one came to listen for it as one listens for the next yap when a dog is baying the moon. It was an unspeakably repellent sound; a foul bubbling and retching, as though the man's bowels were being churned up within him. Once when he struck a match I saw that he was a very old man, with a grey, sunken face like that of a corpse, and he was wearing his trousers wrapped round his head as a nightcap, a thing which for some reason disgusted me very much. Every time he coughed or the other man swore, a sleepy voice from one of the other beds cried out:

'Shut up! Oh, for Christ's—sake shut up!'

I'll take the Parisian kitchen over that anyway. His description of the old man's cough, in particular, makes me never want to sleep in a room with another person ever again. And Orwell seems to share this notion, because he quickly starts to drift from house to house, even "sleeping" in the outdoors sometimes, though in those cases sleep is actually impossible due to harassment by police. Instead he and the other tramps are forced to sit on benches, nodding in and out of consciousness, never really resting. It's this that sends them into the horrible flophouses. Orwell points out that three or four hours of bad sleep is still a significant improvement over a night spent out on the street.

In general, Orwell's life as a London drifter is far more dreary and depressing than his time in Paris. Largely because he has nothing to do, nowhere to be...and he's slowly starving due to the tramp's diet of "tea-and-two-slices" which he describes as such:

"You discover the boredom which is inseparable from poverty; the times when you have nothing to do and, being underfed, can interest yourself in nothing. For half a day at a time you lie on your bed, feeling like the jeune squelette in Baudelaire's poem. Only food could rouse you. You discover that a man who has gone even a week on bread and margarine is not a man any longer, only a belly with a few accessory organs."

In fact, Orwell rightly spends quite a lot of time describing the impact of hunger on a man's will, and identifies it as perhaps the chief means by which respectable people are debased and turned into aimless, wandering tramps. I find these to be among the most moving passages in the whole book:

"You discover what it is like to be hungry. With bread and margarine in your belly, you go out and look into the shop windows. Everywhere there is food insulting you in huge, wasteful piles; whole dead pigs, baskets of hot loaves, great yellow blocks of butter, strings of sausages, mountains of potatoes, vast Gruyère cheeses like grindstones. A snivelling self-pity comes over you at the sight of so much food. You plan to grab a loaf and run, swallowing it before they catch you; and you refrain, from pure funk."

And earlier, speaking on the consolations of poverty:

For, when you are approaching poverty, you make one discovery which outweighs some of the others. You discover boredom and mean complications and the beginnings of hunger, but you also discover the great redeeming feature of poverty: the fact that it annihilates the future. Within certain limits, it is actually true that the less money you have, the less you worry. When you have a hundred francs in the world you are liable to the most craven panics. When you have only three francs you are quite indifferent; for three francs will feed you till tomorrow, and you cannot think further than that. You are bored, but you are not afraid. You think vaguely, 'I shall be starving in a day or two—shocking, isn't it?' And then the mind wanders to other topics. A bread and margarine diet does, to some extent, provide its own anodyne.

I've experienced this myself during deliberate periods of fasting. Once your body starts burning fat for fuel, and you're no longer subject to the energetic highs and lows that come from a high-glucose diet, a queer sort of emotional stability sets in. I didn't like it at all, precisely because of that 'shocking, isn't it?' attitude that Orwell points out. Not to say that the psychological changes one feels during deliberate fasts approach those experienced during long periods of real starvation. Especially when one considers all the counterintuitive changes in attitude that accompany a real transition into the state of desperate poverty all people, in all times have so feared:

"...there is another feeling that is a great consolation in poverty. I believe everyone who has been hard up has experienced it. It is a feeling of relief, almost of pleasure, at knowing yourself at last genuinely down and out. You have talked so often of going to the dogs—and well, here are the dogs, and you have reached them, and you can stand it. It takes off a lot of anxiety."

And yet that same, dulling hunger pushes Orwell and other London tramps to take charity from religious organizations, where the petty humiliations reach their ultimate climax:

"Uncomfortably we took off our caps and sat down. The lady handed out the tea, and while we ate and drank she moved to and fro, talking benignly. She talked upon religious subjects—about Jesus Christ always having a soft spot for poor rough men like us, and about how quickly the time passed when you were in church, and what a difference it made to a man on the road if he said his prayers regularly. We hated it. We sat against the wall fingering our caps (a tramp feels indecently exposed with his cap off), and turning pink and trying to mumble something when the lady addressed us. There was no doubt that she meant it all kindly. As she came up to one of the north country lads with the plate of buns, she said to him:

'And you, my boy, how long is it since you knelt down and spoke with your Father in Heaven?'

Poor lad, not a word could he utter; but his belly answered for him, with a disgraceful rumbling which it set up at sight of the food. Thereafter he was so overcome with shame that he could scarcely swallow his bun. Only one man managed to answer the lady in her own style, and he was a spry, red-nosed fellow looking like a corporal who had lost his stripe for drunkenness. He could pronounce the words 'the dear Lord Jesus' with less shame than anyone I ever saw. No doubt he had learned the knack in prison."

Orwell and a large group of other tramps later attend a church service where they outnumber the worshippers, and so they indulge in stamping their feet and openly heckling the preacher and congregation:

“It was so different from the ordinary demeanour of tramps—from the abject worm-like gratitude with which they normally accept charity. The explanation, of course, was that we outnumbered the congregation and so were not afraid of them. A man receiving charity practically always hates his benefactor—it is a fixed characteristic of human nature; and, when he has fifty or a hundred others to back him, he will show it.”

Yet we soon meet a man who performs charity in a way the tramps approve of. I suggest you stop right now and try to form a picture of ‘charity done right’ before reading on. It made perfect sense to me in hindsight, and corresponded to instances of unpretentious generosity I’d seen before, but I doubt I would’ve been able to form a clear idea of what it might look like:

“Presently the clergyman appeared and the men ranged themselves in a queue in the order in which they had arrived. The clergyman was a nice, chubby, youngish man, and, curiously enough, very like Charlie, my friend in Paris. He was shy and embarrassed, and did not speak except for a brief good evening; he simply hurried down the line of men, thrusting a ticket upon each, and not waiting to be thanked. The consequence was that, for once, there was genuine gratitude, and everyone said that the clergyman was a—good feller. Someone (in his hearing, I believe) called out: ‘Well, he’ll never be a—bishop!’—this, of course, intended as a warm compliment.”

This in contrast to the state-run generosity of the Spike:

“It appeared from what they said that all spikes are different, each with its peculiar merits and demerits, and it is important to know these when you are on the road. An old hand will tell you the peculiarities of every spike in England, as: at A you are allowed to smoke but there are bugs in the cells; at B the beds are comfortable but the porter is a bully; at C they let you out early in the morning but the tea is undrinkable; at D the officials steal your money if you have any—and so on interminably. There are regular beaten tracks where the spikes are within a day’s march of one another. I was told that the Barnet-St Albans route is the best, and they warned me to steer clear of Billericay and Chelmsford, also Ide Hill in Kent. Chelsea was said to be the most luxurious spike in England; someone, praising it, said that the blankets there were more like prison than the spike. Tramps go far afield in summer, and in winter they circle as much as possible round the large towns, where it is warmer and there is more charity. But they have to keep moving, for you may not enter any one spike, or any two London spikes, more than once in a month, on pain of being confined for a week.”

Basically, spikes are homeless shelters spread out all over England, wherein tramps can stay and receive a little food, a place to sleep, and an absolutely revolting bath:

“The scene in the bathroom was extraordinarily repulsive. Fifty dirty, stark-naked men elbowing each other in a room twenty feet square, with only two bathtubs and two slimy roller towels between them all. I shall never forget the reek of dirty feet. Less than half the tramps actually bathed (I heard them saying that hot water is ‘weakening’ to the

system), but they all washed their faces and feet, and the horrid greasy little clouts known as toe-rags which they bind round their toes. Fresh water was only allowed for men who were having a complete bath, so many men had to bathe in water where others had washed their feet. The porter shoved us to and fro, giving the rough side of his tongue when anyone wasted time. When my turn came for the bath, I asked if I might swill out the tub, which was streaked with dirt, before using it. He answered simply, 'Shut yer—mouth and get on with yer bath!' That set the social tone of the place, and I did not speak again.”

And this, I think, is why Orwell is about as popular with the right as with the left: he takes an almost libertarian pleasure in describing (and in the case of *1984*, inventing) the most grotesque forms of governmental overreach and mission failure imaginable. And basically no matter where you stand politically, it's hard not to enjoy that kind of stuff.. No horrific detail escapes him. You can see his future as a great writer of dystopian fiction in his depiction of the Spike's dining-room:

“In the morning, after breakfast and the doctor's inspection, the Tramp Major herded us all into the dining-room and locked the door upon us. It was a limewashed, stone-floored room, unutterably dreary, with its furniture of deal boards and benches, and its prison smell. The barred windows were too high to look out of, and there were no ornaments save a clock and a copy of the workhouse rules. Packed elbow to elbow on the benches, we were bored already, though it was barely eight in the morning. There was nothing to do, nothing to talk about, not even room to move. The sole consolation was that one could smoke, for smoking was connived at so long as one was not caught in the act. Scotty, a little hairy tramp with a bastard accent sired by Cockney out of Glasgow, was tobaccoless, his tin of cigarette ends having fallen out of his boot during the search and been impounded. I stood him the makings of a cigarette. We smoked furtively, thrusting our cigarettes into our pockets, like schoolboys, when we heard the Tramp Major coming.

Most of the tramps spent ten continuous hours in this comfortless, soulless room. Heaven knows how they put up with it. I was luckier than the others, for at ten o'clock the Tramp Major told off a few men for odd jobs, and he picked me out to help in the workhouse kitchen, the most coveted job of all. This, like the clean towel, was a charm worked by the word 'gentleman'.

At three I went back to the spike. The tramps had been sitting there since eight, with hardly room to move an elbow, and they were now half mad with boredom. Even smoking was at an end, for a tramp's tobacco is picked-up cigarette ends, and he starves if he is more than a few hours away from the pavement. Most of the men were too bored even to talk; they just sat packed on the benches, staring at nothing, their scrubby faces split in two by enormous yawns. The room stank of ennui.”

I couldn't find any pictures of a Spike's interior, but I expect one of a workhouse will



serve. Just imagine this, but with worse decor and food:

When Orwell tries to speak to one of his companions(also a tramp)of improvements that might be made, the fear of the mob rears its ugly head once more:

"...I told him about the wastage of food in the workhouse kitchen, and what I thought of it. And at that he changed his tone instantly. I saw that I had awakened the pew-renter who sleeps in every English workman. Though he had been famished along with the others, he at once saw reasons why the food should have been thrown away rather than given to the tramps. He admonished me quite severely.

'They have to do it,' he said. 'If they made these places too comfortable, you'd have all the scum of the country flocking into them. It's only the bad food as keeps all that scum away. These here tramps are too lazy to work, that's all that's wrong with them. You don't want to go encouraging of them. They're scum.'

I produced arguments to prove him wrong, but he would not listen. He kept repeating:

'You don't want to have any pity on these here tramps—scum, they are. You don't want to judge them by the same standards as men like you and me. They're scum, just scum.'

It was interesting to see the subtle way in which he disassociated himself from 'these here tramps'. He had been on the road six months, but in the sight of God, he seemed to imply, he was not a tramp. I imagine there are quite a lot of tramps who thank God they are not tramps."

This causes Orwell to wonder at the reasons for the fall of all these once respectable, wandering people. He gives several, along with reasons for their entrapment within such a wretched life:

*"It is a curious thing, but very few people know what makes a tramp take to the road. And, because of the belief in the tramp-monster, the most fantastic reasons are suggested. It is said, for instance, that tramps tramp to avoid work, to beg more easily, to seek opportunities for crime, even—least probable of reasons—because they like tramping. I have even read in a book of criminology that the tramp is an atavism, a throw-back to the nomadic stage of humanity. And meanwhile the quite obvious cause of vagrancy is staring one in the face. Of course a tramp is not a nomadic atavism—one might as well say that a commercial traveller is an atavism. **A tramp tramps, not because he likes it, but for the same reason as a car keeps to the left; because there happens to be a law compelling him to do so. A destitute man, if he is not supported by the parish, can only get relief at the casual wards, and as each casual ward will only admit him for one night, he is automatically kept moving. He is a vagrant because, in the state of the law, it is that or starve. But people have been brought up to believe in the tramp-monster, and so they prefer to think that there must be some more or less villainous motive for tramping."***

This seems reductive on its face, but Orwell isn't talking about poverty as a whole here, but rather the particular English phenomenon wherein the impoverished are basically on a never-ending walkabout around the countryside. As for the actual cause of poverty, Orwell is anything but reductive. Again and again he states that these are normal men, whom due to an accident, a crime, or one poor decision, have been reduced to tramping. He draws a distinction between what he observes and the American 'hobo' culture that was prevalent at this time:

"Deliberate, cynical parasitism, such as one reads of in Jack London's books on American tramping, is not in the English character. The English are a conscience-ridden race, with a strong sense of the sinfulness of poverty. One cannot imagine the average Englishman deliberately turning parasite, and this national character does not necessarily change because a man is thrown out of work. Indeed, if one remembers that a tramp is only an Englishman out of work, forced by law to live as a vagabond, then the tramp-monster vanishes. I am not saying, of course, that most tramps are ideal characters; I am only saying that they are ordinary human beings, and that if they are worse than other people it is the result and not the cause of their way of life."

I have little doubt that this was true in Orwell's time, but I think of all his assertions about the nature of poverty and homelessness, this one is the most dated. In Orwell's time the 'tramp culture' was not a drug culture. At least in the United States, it now most certainly is. The National Coalition for the Homeless reports that "38% of homeless people are alcohol dependant, and 26% are dependent on other harmful chemicals."¹² And how does Orwell describe his fellow tramps, in terms of substance use?

"...take the idea that all tramps are drunkards—an idea ridiculous on the face of it. No doubt many tramps would drink if they got the chance, but in the nature of things they cannot get the chance. At this moment a pale watery stuff called beer is sevenpence a pint in England. To be drunk on it would cost at least half a crown, and a man who can command half a crown at all often is not a tramp."

No other drugs are even mentioned. But that isn't even the most shocking absence from Orwell's picture of poverty. What struck me first about Orwell's companions is that despite all of them being unemployed, half-starving wanderers without any social ties... none of them seem to be (extremely)mentally ill. Orwell never once suggests that these men are homeless because of any mental deficiency or disorder. He always maintains that they are just like people he knew in his 'respectable life', but poorer. Some are eccentric in the extreme, but nothing like many of the homeless one now sees on the streets of American cities.

This is likely because in 1808, English parliament authorized every county to build its own asylum, and in 1845 it became compulsory for the counties to do so. And "by the end of the century there were as many as 120 new asylums in England and Wales, housing more than 100,000 people."¹³ Now as to the conditions within these asylums, please consult someone who knows more about the history of mental illness than I do. Maybe you could start with Scott's review of Madness and Civilization. But it at least seems a likely explanation for the utter and complete absence of the (extremely)mentally ill from Orwell's Paris and London. Orwell never even mentions the problem. Likely it was a given to him that 'tramps' were mostly sane, able-bodied men, and that anyone that was truly insane would've been immediately locked up. I can't be sure, because he never mentions mental illness or the asylum system at all. It's as if it's not even worth mentioning.

After a brief discussion of the sexual politics of trampdom(men out number women about ten to one, and all tramps are celibate unless they are lucky enough to come into the money needed for the cheapest prostitutes). Orwell asserts that sexual starvation is every bit as degrading to the tramps as physical starvation. I find it fascinating that he fixates on celibacy as an indignity reserved for "cripples and imbeciles" whereas now it is an indignity reserved for...lots and lots of people across each and every social class. No doubt people from Orwell's time would marvel at the modern paradox, wherein sexual liberation has resulted in unprecedented rates(and still growing!) of celibacy in

¹²<https://www.addictioncenter.com/addiction/homelessness/>

¹³<https://historicengland.org.uk/research/inclusive-heritage/disability-history/1832-1914/the-growth-of-the-asylum/>

prosperous countries like the US and Japan. And there seems to be no doubt in his mind that the tramps' lack of sexual and romantic access plays a significant role in their growing alienation from respectable society as a whole. Huh. I think this aspect of homelessness probably isn't discussed enough. In contrast to our image of highly social hobos huddled around campfires, there's likely a large portion of homeless who live like hikkimoris without a room.

Finally we come to Orwell's proposed improvements to the system. And to his credit, the proposed changes are concrete, modest, and presumably attainable:

“Granting the futility of a tramp's life, the question is whether anything could be done to improve it. Obviously it would be possible, for instance, to make the casual wards a little more habitable, and this is actually being done in some cases. During the last year some of the casual wards have been improved—beyond recognition, if the accounts are true—and there is talk of doing the same to all of them. But this does not go to the heart of the problem. The problem is how to turn the tramp from a bored, half alive vagrant into a self-respecting human being. A mere increase of comfort cannot do this. Even if the casual wards became positively luxurious (they never will) a tramp's life would still be wasted. He would still be a pauper, cut off from marriage and home life, and a dead loss to the community. What is needed is to depauperize him, and this can only be done by finding him work—not work for the sake of working, but work of which he can enjoy the benefit. At present, in the great majority of casual wards, tramps do no work whatever. At one time they were made to break stones for their food, but this was stopped when they had broken enough stone for years ahead and put the stone-breakers out of work. Nowadays they are kept idle, because there is seemingly nothing for them to do. Yet there is a fairly obvious way of making them useful, namely this: Each workhouse could run a small farm, or at least a kitchen garden, and every able-bodied tramp who presented himself could be made to do a sound day's work. The produce of the farm or garden could be used for feeding the tramps, and at the worst it would be better than the filthy diet of bread and margarine and tea. Of course, the casual wards could never be quite self-supporting, but they could go a long way towards it, and the rates would probably benefit in the long run. It must be remembered that under the present system tramps are as dead a loss to the country as they could possibly be, for they do not only do no work, but they live on a diet that is bound to undermine their health; the system, therefore, loses lives as well as money. A scheme which fed them decently, and made them produce at least a part of their own food, would be worth trying.”*

Equally modest is the book's brief conclusion:

“My story ends here. It is a fairly trivial story, and I can only hope that it has been interesting in the same way as a travel diary is interesting. I can at least say, Here is the world that awaits you if you are ever penniless. Some days I want to explore that world more thoroughly. I should like to know people like Mario and Paddy and Bill the moocher, not from casual encounters, but intimately; I should like to understand what

really goes on in the souls of plongeurs and tramps and Embankment sleepers. At present I do not feel that I have seen more than the fringe of poverty.

Still I can point to one or two things I have definitely learned by being hard up. I shall never again think that all tramps are drunken scoundrels, nor expect a beggar to be grateful when I give him a penny, nor be surprised if men out of work lack energy, nor subscribe to the Salvation Army, nor pawn my clothes, nor refuse a handbill, nor enjoy a meal at a smart restaurant. That is a beginning.”

But unfortunately, though this might have been a beginning for Orwell, it wasn't much of a beginning for the rest of the literary world. Though careful studies of homelessness have been made, there's no other popular book that attempts to do what Orwell did, unless you count stuff like *On the Road*, which I think obscures more than illuminates what a life in poverty is actually like. I can't fully convey how much I think a book like *Down and Out in San Francisco, Portland, Honolulu and Every other American City that has a Homelessness Crisis* needs to be written, because I think our collective attitude towards homelessness isn't so different from the one Orwell notices in his middle-to-upper-class fellows. It is something that everyone knows about as a general phenomenon, but that almost no one knows about in its particulars. We tend to regard it as something inevitable, unchangeable, or at least, unchangeable by *us*. Like litter and terrible traffic. But then, there are plenty of places on earth that don't have much litter *or* traffic.

I wish the chapter on life in the Spike, in particular, could be read by every politician creating legislation that attempts to combat homelessness. At the very least, it might serve as a cautionary tale for anyone trying to create a shelter system. Beyond that, I wish this and all the rest of Orwell's excellent non-fiction wasn't quite so overshadowed by *1984* and *Animal Farm*. There's so much here that even a thirteen thousand word review can't hope to cover.

As for Orwell's own proposed improvements to the Spikes...I can at least say they are improvements over locking hundreds of hungry men in an empty room for ten hours. Perhaps they are even good improvements, but I find myself uncomfortable with turning these homeless shelters into more workhouses. Ironically enough, workhouses were formally abolished by a law passed the very year Orwell probably came up with these ideas. But he at least gestures at the main failings of the system; chiefly that they are by their very nature nearly inescapable because they enforce idleness, and treat men who are desperate for work like parasites, instead of as men who are desperate to work.

The chapters on Paris make no mention of any welfare system that Orwell might've drawn upon, and instead stress the desperate search for work and food as absolutely paramount for survival. I can't find any evidence for a robust welfare system in 1920s Paris, outside of a few laws passed forcing employers to provide insurance in the case of illness, maternity, etc. In this contrast Orwell provides some potent ammunition for opponents of welfare; for on the one hand we have Orwell toiling away in a kitchen but otherwise living a vibrant, interesting life, and on the other we have him drifting from one

Kafkaesque state-run welfare prison to the next in order to survive. What's great about Orwell is that he sees no real dichotomy here; to him these are both symptoms of the same problem, that being upper-class snobbery and fear of the mob, along with an unwillingness to take the suffering and toil of poor people as a real impetus for change, whether it be top-down or bottom-up. He views these problems as stemming from tractable moral failings at the individual level. And though he goes a long way in making the classism of his time real and palpable to we moderns, I still don't think we can really understand it, in the same way we can't really understand a number like a googolplex. We have no means of understanding it, because classism today is so obscured by its manifestations, whereas in his time you simply were part of a class, deep in your soul, and were treated accordingly. Appearances and even habits were divorced from your class, which was something invisible, bound to you from birth, like a ghost or a horcrux or something. Take this interaction Orwell has with a soldier-turned-Spike-overseer:

"So you are a journalist?"

'Yes, sir,' I said, quaking. A few questions would betray the fact that I had been lying, which might mean prison. But the Tramp Major only looked me up and down and said:

'Then you are a gentleman?'

'I suppose so.'

He gave me another long look. 'Well, that's bloody bad luck, guv'nor,' he said; 'bloody bad luck that is.' And thereafter he treated me with unfair favouritism, and even with a kind of deference. He did not search me, and in the bathroom he actually gave me a clean towel to myself—an unheard-of luxury. So powerful is the word 'gentleman' in an old soldier's ear."

The notion that Orwell might be lying never occurs to the major. The fact that Orwell is now a tramp like all the others doesn't matter either. What matters is that he was a gentleman, and therefore *still is* a gentleman, deep down in chakras. I suppose this is the cultural groundwork for the income-independent classism discussed at length in Scott's review of *Fussell on Class*. I imagine Orwell was laughing at himself on the inside, disappointed in the knowledge that even months of starving and working as a scullion couldn't change the fact that he was an upper middle class Etonian that served in the imperial police. But of course it's that tension that makes this and all the rest of Orwell's non-fiction so interesting. Whether he's taking down a stampeding Burmese elephant in *Shooting an Elephant* or fighting Franco's fascists alongside anarcho-syndicalists in *Homage to Catalonia*, there's always a sense that he's somewhere he's not supposed to be, bringing back forbidden knowledge from unexplored moral territory, so that it might sit comfortably on middle-class and public school library bookshelves. Orwell's genius, as I see it, is in *not being a genius*. He was merely among the first to realize that ugly, uncouth, and unconscionable places and people might be worth a closer look, and that the lives of such people had much broader political and social

significance than the reading public had yet dared to imagine. If nothing else, *Down and Out in Paris and London* should serve as inspiration to journalists and writers everywhere; it's proof that if one wishes to write an important book, one need only write truthfully about the vaguely terrifying parts of society that the average person often sees, but never enters.

Enlightenment Now by Stephen Pinker

Energy Enlightenment and the Better Angels of our Exotherm

When I was a kid, my mother always had a shelf of serious popular science books, and among the authors represented was Stephen Pinker, a Harvard professor and linguist. I'm not the language zealot that she is, but I kept track of Pinker, who has come to prominence in the last 10 years for a pair of books about violence and the well-being of humanity. In the first, (*The Better Angels of our Nature*) he argues – with reams of data – that violence has declined drastically in modern times, and explores some causal themes. In the second (*Enlightenment Now*) he further demonstrates a remarkable, consistent, and progressive improvement in the material, social, and intellectual well-being of humanity, and argues that that the primary cause of this improvement is the rise of Enlightenment values, including science and humanism, which took root in Europe in the 1700s.

I am basically convinced with respect to the decline of violence and improvement in quality of life. Two thirds of *Enlightenment Now* consists of a train of short chapters presenting data to show how longevity, health, wealth, knowledge, freedom, crime, safety, happiness, etc. have improved drastically in the last two hundred years. It's not a subtle book, and its triumphal tone is an odd fit to the mood of our times, but there's a ton of evidence to support the argument that if one was forced to choose a time and place to be dumped – [Rawls-fashion](#) – into the world, 'right about now' would be a pretty good choice. Pinker does grapple with the existential risks of climate change and nuclear war, and acknowledges they are real. Unsurprisingly, he argues that reason, science, and humanism are our best tools for overcoming them, and I agree.

I am less convinced of the Enlightenment as the cause of this dramatic improvement. I am not a historian, but my amateur sense is that there have been a lot of smart people working out principles of philosophy, logic, and the intricacies of the natural world for at least a few thousand years. But something else started in the 1700s, accelerated sharply in the 1800s, and then exploded globally in the 20th century: the development of techniques to burn fossil fuel to liberate immense quantities of energy. This suddenly enabled humans to perform useful tasks at superhuman scale, and I believe it is a much

more powerful force than the achievements of any cohort of philosophers. If this is true, it has serious implications for the future of the benign trends that Pinker celebrates.

To understand the force of this argument, the reader will need a quantitative sense of energy at the human (and superhuman) scale. This is an essay that I have been meaning to write for some time, both because it ties together many of the themes that captivate my personal and professional interest, and because I believe **the average citizen doesn't understand how profoundly energy fuels and enables every aspect of life**, both primitive and modern.

In simple terms, energy is a property that provides the ability to do work. Work has a specific technical meaning, but for practical purposes it means roughly what blue-collar people think it means – for example, energy must be provided to do the work of hauling water from a well, pushing a vehicle along a road against the resisting force of aerodynamic drag, or driving a flow of electric current through a resistance to create light. Energy comes in a number of forms (kinetic, thermal, chemical, potential, etc.), and humans use it both to do physical work and to perform chemical and industrial processes, heat or cool buildings, cook food, etc. This diversity of uses reflects the fundamental importance of energy, which extends to our physical bodies – like all organisms, we require energy to survive. Food is the fuel that allows our bodies to do work, and without it we quickly die.

Energy is universal and quantitative. Universal because it cannot be created or destroyed, and because its various forms can be interconverted, subject to natural laws and practical limitations. Quantitative because it can be measured, and certain tasks absolutely require a defined amount of it. If it requires 10 units of energy to get my car to the top of the hill, and my battery or tank only contains 8, the car will predictably stop short of the summit.

The proper scientific unit of measure for energy is the Joule (J), which is a tiny amount – about as much as is released when a sandwich falls off a table and hits the floor. An iPhone 5s stores about 20,000 J of electrical energy, an Oreo™ cookie contains about 300,000 J of food energy, and a gallon of gas releases about 120,000,000 J of thermal energy when it burns. Because the joule is such a tiny amount, we have other practical units of energy that civilians are more familiar with, including the kilowatt-hour (kWh), which is equal to 3.6 million Joules.

Is a kWh a large amount of energy, or a small amount? The fascinating answer is: both, and this starts to get at the point I'm trying to make.

On one hand, it's a relatively piddling amount in modern terms, equivalent to the thermal energy in a few tablespoons of gasoline. In less than an hour I can tap a kWh effortlessly from the outlet under my desk, and the most amazing thing is that Central Maine Power will only charge me fifteen cents for it.

On the other hand, on the scale of a human body, one kilowatt-hour is a formidable quantity. Imagine pushing a car up a steep grade for over a mile – that's a kWh. I could pedal a bicycle coupled to a grain grinder or other machinery all day and struggle to deliver a single kWh worth of energy.

To bulk up our intuition about energy at the human scale, it's helpful to understand a related concept, *power*. While it is common in the civilian world to mix up energy and power, the concepts are related but distinct in an important way. Specifically, power (in the engineering sense), is simply the rate at which energy is delivered. If one joule is delivered per second, this is described as a 1 watt flow of power. So an old-fashioned 100W lightbulb consumes 100J of electric power per second, most of which is wasted as heat; a modern LED bulb might deliver the same amount of light while consuming only 15 J per second. If the old-fashioned bulb is operated for one hour (3600 seconds), in total it will use 360,000 Joules, or 0.1kWh.

It turns out that if you ask the average healthy non-athlete to pedal a bicycle (or climb a ladder, or some other efficient means of producing power at a sustained pace), you find that a human body can only deliver useful work at a rate of about 100W over a period of hours, and significantly less on average, since we require hours of rest and sleep. **And for hundreds of thousands of years, that was pretty much all the energy we had.** The Bible says "In the sweat of thy face shalt thou eat bread, till thou return unto the ground" and back in the day that was pretty much the size of it. The great majority of people foraged or toiled in fields to grow crops, and they did it pretty much their entire lives.

Naturally the proximal source of that energy was the food they ate, but its ultimate source was the sun, which powered the photosynthesis that stockpiled that energy in the crops and livestock in the form of sugars, fats, starches, and the like. This was a

serious limitation, because photosynthesis is relatively inefficient at turning sunlight into stored energy. According to Wikipedia, [typical crops are only about 1% efficient](#) in turning the sunlight that strikes a field or forest into biomass, so it takes a lot of land (or a lot of time) to produce a given amount of usable plant energy. In many climates (e.g. deserts) the conversion is many orders of magnitude less efficient.

As a result, for millennia our ancestors were fundamentally limited by the strength of their bodies and the relatively modest efficiency with which crops could turn sunlight into food and fuel. What about beasts of burden? The more fortunate among our ancestors had access to an ox or perhaps a horse, which can deliver a modest multiple on the power of the human body. But like humans, draft animals were solar-powered, and their calorie needs were likewise multiplied – a horse or cow required the output of several acres of land for its fodder, and this land could not be used to grow food for humans.

Of course beasts can be eaten as well as worked, but here again, the amount of land required to feed a person on meat is far more than the cropland required to feed them directly on plant-based foods, which is why meat was (and probably still should be) considered a luxury.

Worse, the most productive staple crops require cooking (more energy) to be readily digestible, and cooking was likewise done using wood, which required still more land. Firewood supply was limited in the more populated areas – google ‘coppicing’ or ‘pollarding’ to get a sense for how the supply of precious renewably-grown combustibles was husbanded in those times. Using land to grow fuelwood traded off against using the same land to grow food crops.

If energy is universal and quantitative, and energy for humans comes in the form of food, it should be possible to relate the amount of food we eat to the amount of work we can do. The typical human diet contains about 2000 Calories per day; the Calorie is an archaic unit of energy equal to 4180 Joules. So 2000 Calories is about 8,400,000 Joules or 2.3 kWh. To put our diets in Power terms, I am delighted to discover that typing “2000 Calories per day in watts” into Google yields the following:

2000 (kilocalories per day) =

96.8518519 watts

That is, we eat food energy at an average rate of about 100 watts, and this sets an absolute limit on the amount of physical work we can do; in actuality we'd be lucky to deliver 100W of work for 8 hours per day, with the other $\frac{2}{3}$ of the calories given over to the business of living. And because food is fuel, serious endurance athletes need much more – [up to 8000 Calories per day](#).

To sum it up, **in pre-industrial times our ancestors lived 'land to mouth'**. Life went along this way for hundreds of thousands of years, and though it changed in appearance and intensity with the invention of agriculture, the same fundamental limitations were in place. At best, people carefully husbanded a limited 'working capital' of stored foods, livestock, and standing timber; however, despite primitive tools it was all too easy to over-exploit the productive ecological base and get in an ugly situation, as [Jared Diamond details](#) in cases including Easter Island, Greenland, and others. Life was nasty, brutish, and short in the myriad ways described in the 'before' section of Pinker's books.

But things started to change in a serious way when people discovered that they could tap ancient energy reservoirs of stored sunlight. For a fascinating early example, I recommend an online article called "[Medieval Smokestacks: fossil fuels in pre-industrial times](#)", on the subject of peat as an energy source. Peat is the remnants of plant matter that accumulates over millennia in wetland areas, protected from decay by the lack of oxygen – this is actually the first step in the much longer process that forms coal. Peat can be cut, dried, and burned to liberate thermal energy, and the author, Kris de Decker explores in detail how unique circumstances enabled the people of the area that is now the Netherlands to mine and burn massive prehistoric reserves of it, and thus to liberate themselves from the limitations of their annual allotment of sunlight. The Dutch also mastered the craft of building windmills, which provided mechanical energy to complement the thermal energy from the peat. As a result, they were able to power an impressive array of proto-industrial activity, including glass, brick, ceramics, ships, sugar, salt, soap, spirits, and textiles.

The ability to mine and burn fossilized plants changed the game for the inhabitants of the Low Countries in a material way. By the 1600s, the per-capita annual consumption of peat amounted to about 16 gigajoules per person per year, or about 500W of continuous thermal power, compared with the 50W or less of labor they could manage

on average from their own bodies. And soon this region became far wealthier than neighboring regions, with 60% urbanization compared to the 10% urbanization of the surrounding areas less favorably endowed with peat. Sadly, the peat reserves were eventually depleted, and this combined with competitive coal-fired industrial production from the UK knocked the Netherlands from their perch – by 1820 the country was down to 38% urban population.

Meanwhile, across the English Channel, the real fossil-fired revolution was spinning up. Natural deposits of coal had been in limited regional use for hundreds or thousands of years for metalworking and local heating in coal-bearing regions. But starting around 1700, a sequence of tinkerers, blacksmiths, and engineers invented and refined the steam engine – a machine that used energy liberated by burning fuel to create hot, pressurized steam. That steam could be used to do work – initially to pump water, which was of great value in draining mineshafts and enabling more coal and other minerals to be extracted. But by around 1780 the engines were coupled to flywheels and rotary shafts to drive mechanized equipment that had previously been confined to locations with available water power. The most prominent steam engine inventor was James Watt, who produced a uniquely efficient engine; the scientific unit of power was appropriately named for him. (Watt also devised the unit ‘horsepower,’ equal to 746W, as a product rating tool. He sandbagged a bit so his customers wouldn’t be disappointed; the average horse could deliver somewhat less than 1hp on an ongoing basis).

To say that the invention of the coal-fueled steam engine was a runaway success is a vast understatement. By tapping an immense store of fossilized sunlight, it removed the limitations of plant-fueled musclepower and the vagaries of wind and water power, and catalyzed a chain reaction of growth, wealth, and innovation. Pumping water from mines greatly increased the availability of fossil fuel and minerals. Engines ran blowers for blast furnaces, rolling mills, and a blossoming array of machinery that advanced manufacturing on every axis. In the early 1800s steam engines were adapted to power ships and to transform the rudimentary railways used in mining operations, making fast, convenient transport of people and goods possible. In the following century, convenient liquid petroleum fuels replaced coal, compact internal combustion replaced bulkier steam engines, and mechanization spread to agriculture, with displaced farm workers taking jobs in manufacturing. An immense fossil-powered chemical industry sprang up,

devising among other miracles [the Promethean ability to turn air and water into nitrogen fertilizer](#), solving a major problem in agriculture. And steam power found new life in giant turbines used to generate electricity, literally bringing light and entirely new axes of wealth and convenience – and eventually the information technology that allows me to write and publish this post.

The power that fossil fuels deliver is amazing in both qualitative and quantitative terms. For the reasons described above, in medieval times the average person's access to mechanical power averaged scarcely 100W from the combined efforts of humans, beasts, and a scattering of weak water-powered mills, and perhaps a couple hundred watts of carefully-husbanded firewood. (see [discussion](#)). By the dawn of the enlightenment, the leading economy of Europe had access to an average 500W of thermal power per capita from burning peat alone. By 1900, citizens of the UK consumed on average over 2500W from burning coal alone. And in 2016 the average American consumed a whopping 10,000W of primary energy continuously. This continual torrent of energy enables the amazing material abundance and variety that most of us enjoy, and the everyday superhuman miracles of modern life: I wrote the first draft of this essay in an airplane seven miles above the surface of the earth, blasting effortlessly across the continent at nearly the speed of a thunderclap.

Is the amazing global surge in quality of life primarily due to philosophical advances, or is it primarily the result of discovering a singular lode of stored energy? It's not that enlightenment values are irrelevant to the amazing advances in quality and quantity of life that humans have enjoyed over the last 200 years. I am a huge fan of science, reason, and humanism, and I'm convinced that they have contributed in a central way to the technological progression outlined above – although it seems that early on a surprising number of advances were made by trial and error rather than systematic study and application of scientific principles. But any discussion of improvements in quality of life over this period that doesn't recognize the immense increase in available per-capita energy that fueled and enabled those advances is missing a critical insight.

I think the answer to this question really matters. If Enlightenment philosophy really is the driving force, then it could be reasonable to expect that challenges around the sustainability and environmental impact of burning fossil fuels will look like minor

matters when viewed from the future. In that case, energy historians of the future will conclude that while we used these fuels because they were available and convenient, had they not been there, we would have readily developed other sources of energy nearly as good, and industrial civilization would have developed more or less at the same pace. According to this view, fossil fuel depletion and malign climatic influence are technocratic issues that can be expected to sort themselves out in due course. There may be some minor changes related to the transition to other sources of energy, but the transition can be expected to happen naturally as a result of market forces, and doesn't pose a fundamental danger to the modern quality of life.

But if, on the other hand, the quality-of-life advances are primarily the result of massive increases in per-capita availability of useful energy, then there is a real danger that the peace, prosperity, and broad-based human flourishing of the last 200 years are highly contingent results of a temporary windfall. If so, their depletion could easily reverse those advances – just as the black rock desert goes back to the lizards and ants after the Burning Man festival. If benign progressive trends are primarily a result of a one-time windfall, a bonanza of nearly-free energy unleashed over the last 200 years, then an unwind over a similar span of time is likely to be less than congenial to those who think the arc of history bends inevitably toward justice. If it taps out significantly faster, then all bets are off. Archaeologists point us to civilizations that have fallen; elaborate complex cultures that have disbanded, with their advanced knowledge lost to the nomads who camp in the ruins.

It doesn't take much of a disruption of the material and economic flows of modern life to deflate the progressive instincts, long-term thinking, and warm-hearted embrace of diversity that Pinker celebrates in his book. The 2008 financial crisis was mild by historic standards, but it severely blunted the flow of capital toward forward-looking clean technologies, and unleashed an ugly undercurrent of intolerance in the body public. For those of us working in the clean energy industry this was strikingly clear, with strong popular and investor interest washed away in a torrent of underwater mortgages and 'pocketbook issues'.

The previous, more severe economic crisis of the 1930s came in an energetic time of plenty, yet it concluded in a global nightmare of genocide that ended in a nuclear arms race. If fossil fuel depletion starts to bite faster than clean technologies can comfortably

replace them, or if the global impact of carbon emissions relentlessly drives millions of refugees from major coastal cities, I have a hard time believing that the advances Pinker credits to enlightened principles will be secure.

If this is the path we are on, then successfully executing a rapid, global transition to clean sources of energy is of supreme importance. The growth of solar, wind, and other scalable clean technologies must continue and accelerate consistently. Energy storage and load shifting/management must both advance without a hiccup, and the electrification of transport must displace fossil fuel as quickly as clean capacity can be added to the grid. Liquid fuels should be reserved to particularly thorny technical challenges like air travel, which may need to be curtailed until significant advances can be made in renewable fuels or the volume- and mass-efficiency of clean energy storage. With political leaders dithering in the face of the greatest civic challenge in generations, it appears to be up to engineers and Swedish highschool students to lead the way to an enlightened future.

Exploring Reality: The Intertwining of Science and Religion by John Polkinghorne

As an atheist and scientist, I read *Exploring Reality: The Intertwining of Science and Religion* out of curiosity about Christian perspectives on science. The author's credentials inspired confidence: a former quantum physicist and current Anglican priest, John Polkinghorne seems like exactly the kind of person who would know a lot about both science and religion.

I was not disappointed. The book was clear, concise, illuminating, and beautifully written. Polkinghorne takes the modern scientific understanding of the world absolutely seriously. His descriptions of physical reality are not only insightful, but awe inspiring, teaching me aspects of the universe that I never understood. Its coverage of theology is equally impressive, ranging from an account of the historical Jesus to an exploration of the Trinity to the nature of evil.

The other aspect of the book I appreciated was its combination of confidence, honesty, and respect for other viewpoints. On one hand, Polkinghorne does not hem or haw about his beliefs. Physics is real, not a social construct. God exists. The Godhead is a Trinity. At the same time, he never mocks or disrespects non-believers, never implies they are immoral, and goes out of his way to acknowledge that followers of other religions have spiritually authentic experiences. Even more impressively, Polkinghorne acknowledges the weaknesses of his positions. For example, he acknowledges that there are many inconsistencies and contradictions between the Gospels, and that the diversity of the world's faiths pose a real challenge to the validity of Christianity. His intellectual honesty is refreshing and adds a lot to the book's value. This is a book that deserves to be reviewed chapter by chapter.

Reality

This short chapter is a defense of realism, and should cause the least objection among both Christian and atheist readers. The point is simple: what science and theology discover about the world may not be absolute truth, but they bear significant resemblance to the truth, and are not mere social constructs. Part of why we'd expect this to be so is the surprising nature of our discoveries:

Far from its behaving like epistemological clay in our pattern-seeking hands, capable of being moulded into any pleasing shape that takes the fancy, the physical world frequently proves highly surprising, resisting our expectations and forcing us to extend, in unanticipated ways, the range of our intellectual understanding. In consequence, the feel of actually doing science is undeniably one of discovery, rather than pleasing construction. Theologians can claim something similar about the encounter with God. Time and time

again human pictures of deity prove to be idols that are shattered under the impact of divine reality.

The Causal Nexus of the World

For me, a physics enthusiast, this was the most interesting chapter. It addresses the problem of causality, deficits in our current physical understanding of causality, and theological implications of these deficits.

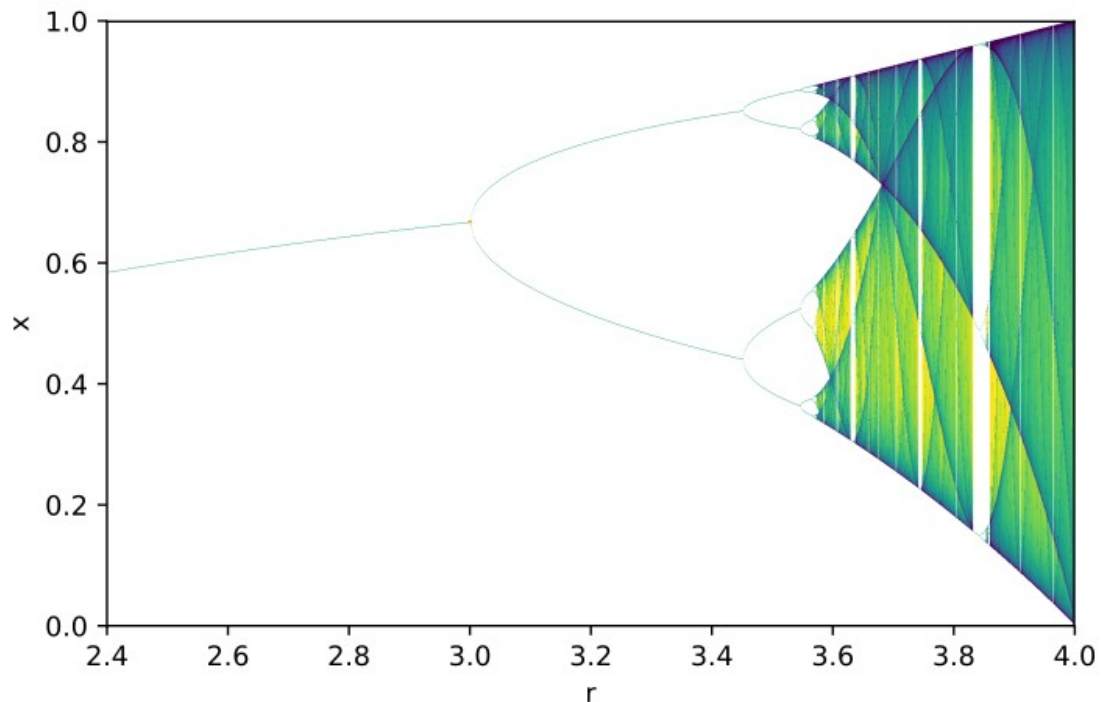
Polkinghorne's basic argument goes like this. Scientists commonly explain the universe using reductionism, which models the whole as the sum of its parts. Hence biology is explained by chemistry, chemistry by physics, and the physics of macroscopic objects by the quantum physics of subatomic particles. But two aspects of the world cast doubt on reductionism. One is [wavefunction collapse](#), which, at least under the Copenhagen interpretation, seems to suggest that macroscopic objects like measurement instruments alter the nature of subatomic particles in a way that other subatomic particles don't—an egregious violation of reductionism. The other is [chaos](#), which makes systems hypersensitive to initial conditions and often causes complex behavior that is very hard to predict with reductionist methods. Worse, these two sources of uncertainty—the inherent randomness of quantum systems and the hypersensitivity to initial conditions of some classical systems—don't play nicely with each other, as they rest on fundamentally contradictory assumptions about the universe. Polkinghorne argues that these difficulties in our understanding of causality open up the possibility that God could intervene in the universe through divine providence. He speculates that in the future, information might become a first-class citizen in physics. This information would have to be incorporated in a way that reflects its semantic as well as syntactic content. "Active information" would then be a means of top-down causality that rescues human agency: human minds, which produce "active information", would have a means of affecting the physical world instead of merely being affected by it. God, too, could use active information as a means of divine providence, "in a manner that operates non-interventionally within the grain of nature, rather than interventionally against it".

There is a lot to unpack in this chapter. As far as I can tell, all of the physics he describes is well accepted and uncontroversial—unsurprising given his previous profession—but he presents it in an engaging and accessible way.

The part I found most interesting, and that I learnt the most from, was Polkinghorne's descriptions of chaos. A chaotic system, among other things, is exponentially sensitive to initial conditions. The smallest perturbation results in large qualitative changes, or to put it more poetically, [the flap of a butterfly's wings in Africa can cause a hurricane in Florida](#). This also implies that the future is unpredictable, even in a perfectly deterministic system, because even the smallest error in measuring its initial conditions (like the error forced upon us by the Heisenberg uncertainty principle) is magnified to momentous proportions in relatively short time.

That's all well and good, but the even more interesting fact is that although chaos seems to complement quantum uncertainty nicely, the two are in fact at tension.

Quantum systems are inherently periodic because Schrodinger's equation describes the time evolution of a discrete set of frequencies, whereas one defining characteristic of chaos is that it has no periods. Chaotic systems create a fractal pattern in phase space (the conceptual 6-dimensional space that combines position and momentum); quantum systems cannot have fractal phase spaces because the phase space is quantized. In fact, we know that quantum mechanics suppresses chaos. Even for large objects, the suppression timescale can be surprisingly short. For the chaotically tumbling Saturnian moon Hyperion, about the size of New York City, quantum mechanics should suppress chaos in 37 years. The fact that it's still tumbling is a result of quantum decoherence, the interaction with the environment that restores the moon to classical behavior with every impacting dust grain and photon.



Bifurcation diagram of the [logistic map](#), a dead simple equation that exhibits an astounding diversity of complex behavior. The ratio between one bifurcation interval to the next approaches Feigenbaum's first constant: 4.669201609... This constant is universal: it holds not just for the logistic map, but for all maps with a single peak. After the periods bifurcate to infinity, which occurs at $r \sim 3.56995$, chaos kicks in. The chaos subsides within islands of stability (white vertical bars). From 3.56995 to 3.82843 is the [Pomeau-Manneville scenario](#), characterized by mostly periodic behavior with bursts of chaos. Image credit: [Morn](#)

Yet another fascinating fact is that novelty is maximized at the [edge of chaos](#):

Too far on the orderly side of that frontier and things are too rigid for there to be more than a shuffling and rearrangement of already existing entities. Too far on the disorderly side, and things are too haphazard for any novelties to persist. A simple example of this principle is afforded by biological evolution. Without a degree of genetic mutation, life would be frozen into the existing range of forms. Too high a mutation rate, and there would be no quasi-stable species on which natural selection could operate.

...and that “chaos”, in fact, is not entirely chaotic:

In dissipative chaotic systems (those in which friction operates), behaviour soon converges onto an intricate but limited portfolio of possible forms, called a ‘strange attractor’. (‘Attractor’ indicates that motions converge upon it; ‘strange’ refers to the fractal character of its structure in phase space.) In those cases where chaos is generated through a cascade of bifurcating possibilities, there is a remarkable universal pattern in the way this happens, characterised by [a new mathematical constant of fundamental significance](#), discovered by Michael Feigenbaum.

The fact that there are mathematical constants that can describe any aspect of chaos is so counterintuitive as to be astounding. The fact that simple equations like the [logistic map](#) can lead to a rich variety of order and chaos illustrates [emergence](#): the generation of complex order from elegant simplicity.

All in all, this was an excellent chapter that showed me a side of reality which I had heard of, but never deeply appreciated. The only weakness is that the tie-in with theology seemed like a classic “God of the gaps” argument: we don’t yet know what causes wavefunction collapse, and even without quantum randomness chaos makes it hard to predict the future, so it’s possible for God to be intervening in the universe. It is *possible*, of course, but Polkinghorne doesn’t elaborate on why this God of the gaps argument is different from any other.

Human Nature: The Evolutionary Context

Evolution is real, and humans are the result of 4 billion years of evolution. Polkinghorne doesn’t doubt this for a moment; in fact, he calls Darwin a great scientist and a genius. Darwin’s contribution to theology is to force theologians to abandon an exclusive focus on *creatio ex nihilo* and put some emphasis on *creatio continua*, a creation where creatures are “allowed to make themselves”.

He does, however, note that humans are qualitatively different from the animals in seven main respects: sapience, language, reason, creativity, morality, religion, and sin. To the extent that animals have these abilities, they are so limited that even calling them by the same names does a disservice to humanity.

Polkinghorne sees a human as a “psychosomatic unity”: a combination of material and mental in an inseparable package. The soul, or the “real me”, is an information-bearing pattern that changes with learning and experience, and which is not inherently immortal. The resurrection would then consist of God reconstructing the information-bearing pattern in the next world.

Polkinghorne argues that the sheer intellectual capacity of humans is beyond the scope of natural evolution. Maybe humans could have evolved the ability to add single-digit numbers, but of what survival advantage is it to be able to do quantum field theory? He applies similar arguments to morality and aesthetics, saying that neither the radical altruism that induces people to sacrifice themselves for strangers, nor the aesthetic sense that allows one to appreciate Mozart, have any survival advantage. To explain these phenomena, he hypothesizes that survival was replaced was satisfaction:

In these noetic realms of rational skill, moral imperative and aesthetic delight —of encounter with the true, the good and the beautiful—other forces are at work to draw out and enhance distinctive human potentialities. Survival is replaced by something that one may call *satisfaction*, the deep contentment of understanding and the joyful delight that draws on enquirers and elicits the growth of their capacities.

Polkinghorne then says, not surprisingly, that the material, moral, and aesthetic spheres of reality were created by God as a gift to mankind.

This chapter was interesting and thought provoking. Moral realism is a fairly common viewpoint among both believers and non-believers, but I had never before heard of aesthetic realism. Associating beauty with a dimension of reality is definitely a unique perspective, but my own perspective is reflected in the adage “beauty is in the eye of the beholder”.

The reason Polkinghorne provides for arguing evolution could not have led to humanity’s vast intellectual capabilities is a common one, first made by Darwin’s detractors shortly after he published his *Descent of Man*. Insights from artificial intelligence research can help us answer the objection. In recent years, AI has managed to beat mankind’s best at chess, Go, and jeopardy; it has helped mathematicians prove theorems; and it has gone toe to toe with professional players in the real-time strategy game Starcraft II. At the same time, AI struggles mightily to do what even a young child can do effortlessly: identify and track objects, speak in coherent sentences, use common sense reasoning, walk, run. Only with recent advances in AI are some of these problems being solved.

This is known as [Moravec’s paradox](#). Humans, it turns out, are very bad at judging the difficulty of problems. Sensorimotor functions come naturally to us because hundreds of millions of years of evolution have first perfected them, then buried them deep in our subconscious. Reasoning, logic, strategy, and math seem difficult to us because we are biologically bad at them. These capabilities were hastily added by evolution in the relatively recent past, and only through the accumulation of millennia of knowledge by

thousands of our species' best and brightest did we advance from being unable to count past 2 to understanding the origin and evolution of the universe. Moravec's paradox is reflected in the human brain by neuron counts: the cerebellum (responsible for motor control) has 70 billion neurons, while the cerebral cortex (responsible for memory, thought, consciousness) has 15 billion.

The Historical Jesus

Jesus was a crucified criminal who died a nobody in a backwater of the Roman Empire, deserted by his followers. Yet we have all heard of him. Why? Polkinghorne's answer is that he really did perform miracles, he really did resurrect after his death, and he really is God. This chapter tries to provide evidence for these extraordinary claims using textual evidence from the New Testament.

Unlike many other Christian apologists, Polkinghorne spends a lot of the chapter pointing out inconsistencies within the New Testament, even within the same book:

Writers in the ancient world were not as concerned about accuracy of detail as are modern historians. What mattered to the ancients was to get the main point across. It is instructive to compare three separate accounts given in Acts (9:1–9; 22:6–11; 26:12–18) of the conversion of Paul on the road to Damascus. Did his companions hear the voice from heaven, or was it only Paul? Did they too see the light of the heavenly vision? Did all fall to the ground, or only Paul? Was Paul told straight away what he was called to do, or was he told that this would be made clear to him in Damascus? The different accounts answer these questions in different ways, despite all coming from the pen of the same author.

Polkinghorne questions the historical accuracy of large chunks of the Gospels, including all the long speeches in John:

The Johannine discourse has a hauntingly timeless quality to it. It is difficult to think that during his lifetime Jesus could actually have spoken in two quite different ways [the one in the Synoptic Gospels, vs. the one in John]. Many believe that John wrote his gospel at the end of a long life of deep reflection on the inner significance of Jesus, expressing his insights in discourses that he then attributed to Jesus.

Polkinghorne doesn't think these inconsistencies are all that important. The Gospels are not biographies or histories, but theological works intended to convey a religious message. The Gospel writers, despite not having the attention to detail of a modern historian, were nonetheless committed to penning the religious truth as they saw it. As evidence of this, Polkinghorne points out that they included details which must have caused them great discomfort. For example, the purpose of baptism was to wash away sins, yet John baptized the supposedly sinless Jesus. Proper burial was a sacred duty among both Jews and Gentiles, yet Jesus says "let the dead bury his own dead." Jesus was supposedly the messiah, yet while being crucified, he cries out "My God, my God,

why have you forsaken me?” Although Polkinghorne doesn’t name this type of argument, it is commonly called the [criterion of embarrassment](#): if an author includes details that contradict his deeply held beliefs, those details are more believable.

What is it that we can tell about Jesus from secular history alone? Certainly not the exact chronology of his life—for one thing, the Gospels have different chronologies. But the gospels as a whole “carry the impress of a single remarkable figure lying behind most of it” and “seem to indicate a single original mind at work”, even if it is not true that “Jesus’ character and convictions are absolutely transparent to us.” Secular history is also enough to tell us that Jesus’ fame probably came from his reputation as a healer, especially as an exorcist. Polkinghorne is surprisingly skeptical of his healing miracles, suggesting that some of the healing encounters may have been successful due to the psychosomatic influence of a charismatic healer.

Another point stressed in this chapter is that Jesus was probably not put to death for having a different Rabbinical interpretation of the Bible, but for making more extraordinary claims. In his Sermon on the Mount, he amplifies the severity of Mosaic Law—to take one example, not only is adultery banned, but so is looking at a woman lustfully. He does this not by appealing to scripture or tradition, but on his own authority: “You have heard it was said to those of ancient times...But I say to you...” (Matthew 5). This claim to the right to modify divine law would have been highly offensive to most Jews. Even more egregious, from the point of view of the Jewish high priest, was Jesus’ throwing the money changers out of the temple during Passover time, a time when Jerusalem was filled with pilgrims and prone to violent riots. Neither the Jewish high priest nor the Roman governor, Pontius Pilate, could tolerate this violent challenge to the authorities, so Pilate had Jesus executed in the interests of stability and order.

This much secular history can reveal. To justify going further and invoking the supernatural, Polkinghorne points to three major failings of the secular account:

1. Plenty of deluded Palestinian religious leaders were executed, to be forgotten by everyone except historians. Why do we all know about Jesus, an executed criminal who was deserted by his followers?
2. Although the gospels are highly discrepant on the issue of Jesus’ resurrection, there is one similarity: Jesus was hard to recognize, to the extent that Mary Magdalene mistook him for the gardener. This similarity is unlikely to arise by chance, and so must have been historical.
3. The empty tomb stories, which take the same form in all four gospels. The strongest justification for taking these stories seriously is that women play the leading role, yet women were not considered capable of being reliable witnesses in a court of law.

Throughout this chapter, Polkinghorne has been as charitable to skeptics as he can possibly be while remaining a Christian. He not only bends over backwards to acknowledge discrepancies within the New Testament, but admits that secular history cannot prove the veracity of supernatural events such as the mass feedings or the

resurrection, let alone Jesus' identity as God. He deserves enormous credit for his intellectual honesty.

That said, his reasons for going beyond secular history are fairly unconvincing. As Polkinghorne said, Jesus was far from the first Palestinian religious leader to be executed, and he would not be the last. This somewhat dilutes the unlikelihood that one of the motley group would found a great religion, even if 99% are doomed to oblivion.

It is still true, of course, that the overwhelming success of Christianity in the Roman Empire is remarkable. As a passionate Roman history enthusiast, I am intrigued by the rapid Christianization of the Roman Empire and its neighboring barbarian tribes. One factor I would point to is the unique nature of Christianity compared to its neighboring pagan religions. Most pagan religions in the Roman Empire were local, ethnic, polytheistic, and tolerant. The gods of the Greeks were the gods of the Greeks; they might be more or less powerful than the gods of the Britons, but certainly do not claim absolute power over Britain. An ancient Greek traveller could pray to Athena while in Athens, to Tanit while in Carthage, and to Jupiter while in Rome, all while considering himself a pious man. Indeed, the Romans considered themselves the most pious people on Earth, for while others only prayed to their local gods, they prayed to all the gods of their subject peoples. Christianity, by contrast, recognized only a single God, a God with supreme power over the entire globe. Central to Christian belief is the Great Commission: "Go therefore and make disciples of all nations, baptizing them in the name of the Father and of the Son and of the Holy Spirit, teaching them to observe all that I have commanded you" (Matthew 28:16–20). Christian emperors of Rome, far from praying to other gods, instead banned pagan rites, closed pagan temples, and put pagan worshippers to death. Whether or not these actions accord with modern notions of morality, they undoubtedly helped Christianity spread across the Roman Empire, one of the largest, most populous, and most well-connected empires in history. Christianity's spread was remarkable, but it is easier to have a remarkable spread when neighboring religions are not even trying to compete on this front.

On Polkinghorne's second point, I do not have much to say. I do not share his judgment that because the gospels share the motif of the risen Christ being hard to recognize, the story must be historical. Undoubtedly the synoptic gospels and John are not completely independent, but draw on some of the same traditions, some oral and some written, that were floating around 1st century Palestine. Mythological motifs can be surprisingly durable, even across different cultures and many millennia of time. For example, the flood story in the Bible is eerily similar to the one in the Epic of Gilgamesh, and elements of Genesis can be found in the Enuma Elish, yet these commonalities are not evidence that the great flood really happened or that Genesis is a literal account of creation.

On Polkinghorne's third point—which is fairly common among Christian apologists—I think most people today exaggerate the low status of ancient women. Roman and Greek myths are full of both human women and goddesses, as even a cursory reading of Hesiod's *Theogony* or Ovid's *Metamorphoses* would show, and pagans worshipped female deities no less piously than male deities. Athenian plays frequently featured

heroines as the main protagonist, including two plays named after them: *Antigone* and *Lysistrata*. The ruins of Pompeii have frescos showing that women frequently owned taverns, bars, and inns. Julia Felix, a particularly famous businesswoman, was a property magnate who rented out properties for residential and business use. Even the New Testament is evidence that women were hardly always confined to the home: there are at least six people named Mary in the gospels; there's the deaconess Phoebe, the missionary Priscilla, and [many more](#).

There is a way to reconcile the severe legal restrictions on women in the ancient world, and the evidence that women often led active lives in business and religion. The law, both religious and secular, represented an ideal—a society's vision for goodness and righteousness. No society ever lives up to its ideals. Modern society has adopted gender equality as a legal and moral ideal, but we only have to visit a construction site, a mine, a daycare, or a nursing home, and take a cursory glance at the gender ratio of the employees there, to know that gender equality is not a reality. Jewish law might have forbidden women from testifying in court, but that did not mean ancient Jews did not have women that they respected, trusted, and admired. The disciples of their religion's founder would seem to be a good target for respect, trust, and admiration.

Divine Reality: The Trinity

This chapter delves deep into Trinitarian theology. It was the most difficult chapter for me to parse, because I never understood trinitarianism. What does it mean for a single God to be composed of Father, Son, and Holy Ghost? How can the Trinity be Three and One at the same time? I'm a fan of Roman history, where seemingly hair-splitting Christological disputes between Arianism, monophysitism, dyophysitism, orthodoxy, and Nestorianism determined loyalties and alliances, and where believers of one Christology persecuted believers of another with relentless zeal. I just wish I could understand it all.

Polkinghorne tries to explain trinitarianism as he sees it. He does it fairly well. He does not deny its profound weirdness, but compares it to quantum theory. Classical mechanics would find it absurd that anything can be both a particle and a wave, or that an object can be in a superposition between two states. Quantum mechanics is clearly unreasonable to a classical mindset. Yet reality is stubborn, and encounters with physical reality have gradually forced scientists to accept the weirdness of the quantum world. We should not expect theological reality to be any less surprising than physical reality. If trinitarianism seems unreasonable, that just indicates our conception of reasonableness must need to be adjusted in light of the early apostles' experiences with theological reality.

So what is the Trinity, exactly? According to Polkinghorne:

Christian experience testifies to the knowledge of God as the Creator of the world (one might say, the Father 'above us'), as God made known to us in human terms in Jesus Christ (the Son 'alongside us') and as God at work in our hearts and lives (the Spirit 'alongside us').

These are not three distinct entities, but “three different modes of encounter with a single divine reality, three contrasting perspectives on what is essentially a monistic deity.” In fact, all three divine Persons participate in all works of God. The Spirit, for example, is said to have hovered over the waters of chaos (Genesis 1:2), while the Word (also known as Jesus) is the one by whom all things were made (John 1:3). But if this is the case, what exactly is the distinction between the three divine Persons? Here Polkinghorne brings in the idea of [appropriation](#), by which a property that in reality belongs to the entire triune God is ascribed to one Person. For what reason, he does not say.

I still don't understand the Trinity, but I take comfort in the fact that Polkinghorne doesn't either:

One is reminded of Bohr's invocation of complementarity to 'explain' wave/particle duality. He said, ask a wavelike question and you will get a wavelike answer, a particlelike question and you will get a particlelike answer and—fortunately for consistency's sake—you cannot ask both questions at the same time. Essentially, this was a rephrasing of the simple fact that that's the way things are. Real understanding had to await the discovery of quantum field theory. Appropriation looks like a theological suggestion of the fitting way to phrase questions about the nature and content of divine revelation.

So appropriation, like particle-wave duality, is a fudge: a stopgap theory which unsatisfyingly papers over our ignorance of the real underlying phenomena, but which nevertheless illuminates more than it obscures. There is not yet a quantum field theory of Christian theology.

The divine Persons are defined not just by their own realms of action, but by their relations with each other. Namely, the Persons are constantly exchanging love, an exchange called “perichoresis”. The Father is the lover, the Son is the beloved, and the Spirit is the love exchanged. The passion of Christ must be seen as an internal event in the Godhead where “the Son suffers in his love being forsaken by the Father as he dies. The Father suffers in his love the grief of the death of the Son.” God's creation of the world is an expression of His internal relationality: he brought the world into being to be an object of divine love.

This picture of the divine Persons constantly exchanging love is charming. I hope it is true. As Polkinghorne says, however, the analogy between appropriation and wave-particle duality is imperfect because the physicist has been far more successful in his quest to understand reality than the theologian. Perhaps this is because theological reality is much harder to probe, as Polkinghorne says; but the fact remains that there is no quantum field theory of theology. If theologians themselves do not fully understand the nature of the Trinity, I certainly have little hope of doing so.

The Nature of Time: Unfolding Story

Does the future already exist? Or is the universe inherently temporal, with a unique “now” and a future that is just now unfolding? Proponents of the block universe believe the former. Polkinghorne is decidedly not a block universe theorist.

The block universe is motivated by our current understanding of physics. First, the laws of physics have no “now” written into them, and instead treats all moments in time equally. So much the worse for physics, Polkinghorne says, if the laws of physics cannot account for such a fundamental aspect of our experience with reality. Second, the laws of special relativity say that not only does time pass at different rates for different observers, but that two simultaneous events in one reference frame are not simultaneous in others. To this, Polkinghorne says that there is a reference frame that we can conveniently define as the reference frame of the universe, namely that of the cosmic microwave background.

The temporal nature of the universe has theological implications. Not just humans, but God himself, apprehends the world temporally. This is exactly how the Bible portrays him. He is not a static God who reveals himself once, but a temporal God who acts within human history. That is why clear theological development happens even in the Old Testament:

I appeared to Abraham, Isaac and Jacob as God Almighty, but by my name ‘the Lord’ I did not make myself known to them (Exodus 6:2–3)

...and why, in the New Testament, God decisively reveals himself in the life, death, and resurrection of Jesus Christ. This revelatory process not only didn’t finish in Jesus’ lifetime (John 16:12–15), but is still ongoing in the community of the modern church.

The ongoing nature of revelation explains why it took until the Councils of Nicea, Constantinople, and Chalcedon, centuries after the New Testament was written, to elucidate the trinitarian nature of God. It also explains why God is seemingly cruel and genocidal in the Old Testament:

The acts of war and genocide that figure so largely in the annals of the deuteronomistic history (Joshua to 2 Kings) do not have parallels in the pages of the New Testament (except, perhaps, for the highly charged symbolism of Revelation). The God who in the Hebrew Bible sometimes proves very dangerous to encounter (Exodus 4:24–26, 2 Samuel 6:6–8) comes to be recognised as the faithful and loving God and Father of our Lord Jesus Christ. This revisionary process continues beyond biblical times. After eighteen centuries, the Church finally came to realise the repugnance of slavery and to question whether a loving God would exact the punishment of everlasting torture for finite transgressions.

The temporal nature of the world also carries theological implications for the world to come. The material world, the first chapter of creation, must decay away due to the

advance of entropy. The second chapter is the world to come, the “realm of realised love” where the resurrected will have everlasting life. This world will still have its own version of matter (remember that Polkinghorne conceives of the human soul as an information-bearing pattern, so something needs to bear the pattern), but the matter will not be relentlessly ground down by entropy. In this non-entropic world, the pain in this world will be healed, the evil effects of sin will be remedied, everyone will be worshipping all the time, and eternal life will never be boring because we’ll be led deeper and deeper into the mysteries of God.

Polkinghorne never uses this analogy, but I get the impression that he thinks of the Bible in the same way that physicists think of Newton’s *Principia* or Darwin’s *Origin of Species*. These were works that revolutionized our understanding of physics and biology, but nobody would say they are 100% correct, let alone a complete explanation of all of physics or biology. Similarly, the books of the Bible were revolutionary theological works, but they are only a starting point from which much more needs to be done. In this analogy, the various ecumenical councils are like scientific conferences, and theologians are like evolutionary biologists who continue to build on Darwin’s legacy.

I like this way of thinking. I find it infuriating when Christians address the undeniable atrocities of the Old Testament, up to and including genocide, with denial, obscurantism, or justification (“the Amalekites were evil and deserved it...”) It is refreshing to see a committed Christian frankly admit to these atrocities while giving a logical reason why they are there.

I will end on an incidental note, by pointing out how surprised I was to read the last half of this sentence:

After eighteen centuries, the Church finally came to realise the repugnance of slavery and to question whether a loving God would exact the punishment of everlasting torture for finite transgressions.

Is it true that “the Church” (which I take to mean at least a plurality of Christians) no longer believes that hell is everlasting torture? When did this shift happen? Why haven’t I heard about this? I wish Polkinghorne spent more than half a sentence on this, because it’s one of the most intriguing parts of this chapter.

The Spirit and the Faiths

I was burning with curiosity about what Polkinghorne would say in this chapter. A common atheist critique of religion, which I share, is that the extreme diversity of the world’s faiths and the sharp contradictions between them puts the validity of all religions into question.

Polkinghorne doesn’t deny this; in fact, he puts it better than I did:

Ask a suitably qualified person in Rome or Jerusalem, Benares or Kyoto, what matter is made of, and in all four cities you will receive the same reply, 'quarks and gluons'. Ask four people in those four cities what is the nature of ultimate reality, and their answers are likely to be very divergent. Does this not show that religious belief is really just a matter of culturally shaped opinion? I do not think so, but I acknowledge that the challenge presented by the diverse cognitive claims of the religions is one that has to be taken very seriously.

Polkinghorne then discusses the feasibility of interfaith dialogue and the best strategies for engaging in it. Given the depth and complexity of each faith, which require a lifetime of learning to grasp, any kind of religious consensus will be a project for this millennium, not this century. In the meanwhile, the best way to engage in interfaith dialogue is neither to hide one's core beliefs—which, for Polkinghorne, would be the trinitarian nature of God—nor to lead with those core beliefs. The former would be dishonest; the latter would be too threatening. Instead, the best approach is to address side questions, such as how each religion fits scientific discoveries into its worldview, in the hope of achieving some amount of common ground and understanding.

So why are there so many contradicting religions? One possibility would be that Christians are completely right and everyone else is completely wrong. Polkinghorne dismisses this possibility. Instead, he believes the Holy Spirit has revealed the truth in differing degrees to different people. Quoting a sermon from Leo the Great in the 5th century:

When the Holy Spirit filled the Lord's disciples on the day of Pentecost, this was not the first exercise of his role because the patriarchs, prophets, priests, and all the holy persons of previous ages were nourished by the same sanctifying Spirit...although the measure of the gifts was not the same

Polkinghorne doesn't expound on how he knows that Christians, and not Buddhists or Muslims, got the most gifts from the Spirit. Partly that is because this chapter is very short, which in turn is because Polkinghorne did not want to repeat what he said elsewhere. But I've never read his other books, so I was disappointed at the lack of detail. In particular, he leaves the most important question unresolved: if someone's religion is largely decided by what country he lives in (as Polkinghorne admitted it does), why does that not mean religion is a social phenomenon not tied to divine reality?

Evil

If God is all-powerful and all-loving, why is there evil in the world? This question, asked by many people over the ages, is a serious challenge to Christian belief. Polkinghorne says as much:

Of all the difficulties that hold people back from religious belief, the question of the evil and suffering in the world is surely the greatest [...] Not only does

it give considerable pause to the enquirer after theism, but it is also one that remains a perpetual challenge and source of perplexity for those of us who are believers.

The challenge is especially serious for natural evils—pandemics, earthquakes, tsunamis. While moral evils such as cruelty and exploitation are the fault of humans, natural evils seem to be the fault of God.

Christian responses have followed three main strategies. The first is to say that Adam and Eve eating the forbidden fruit is what caused sin to enter the world. This explanation is no longer tenable due to the advance of science, which show not only that the Genesis creation story is not literally true, but that death and destruction are far older than the human species. The second approach is to deny the existence of evil. The horrors of the twentieth century, especially the Holocaust, make this untenable. The third and most common strategy is to argue that evil is the necessary cost of greater good.

This third approach is called the free will defense (for moral evil), or the free process defense (for natural evil). God has allowed creatures to create themselves through evolution, and given humans the freedom to shape their own societies. The necessary consequence of being free to choose the good is that one can choose the bad; the necessary cost of creatures being able to make themselves through evolution is the cruel process of natural selection, which necessarily involves mutation-induced cancer, starvation, and mass extinction. The reason that God did not make a universe where all conscious beings would always freely choose the good is that it is logically impossible.

Polkinghorne realizes that this view puts the existence of heaven into doubt. After all, isn't heaven a place where everyone is free, yet does not sin? Maybe the presence of God is what elicits "a full and free acceptance of the divine will". But if this is so, why doesn't God just unveil himself right now, in this world, so that everyone fully and freely accepts the divine will? Polkinghorne's answer is that the unveiling has to happen in stages, and that a full unveiling right now would undermine human autonomy:

It is only after a free decision has been taken to renounce the illusion of human autonomy and to embrace the reality of heteronomy that the nature of God can progressively begin to be revealed with greater clarity and without forcing the individual.

Polkinghorne's "free will" and "free process" defenses for moral and natural evil respectively are not new or radical ideas, but they seem to suggest a kind of deism, where God put a clockwork universe into motion and let its history play out. This is very different from the personal God that most Christians have in mind when they pray for success in their careers or for deliverance from suffering. This deistic view would appear to take away much of the motivation many Christians have for their belief, and much of the comfort Christians derive from their faith.

While I grant that the “free will” defense for moral evil is somewhat convincing—although an Auschwitz inmate being worked to death might beg to differ—the “free process” defense for natural evil is far less so. Why is it inherently good for inanimate beings like viruses to “make themselves”, thereby causing pandemics that kill tens of millions? The freedom argument may be tenable for moral evils, but is much less convincing when the benefactors of that freedom are inanimate objects and the victims of that freedom are human beings.

Ethical Exploration: Genetics

This chapter is a wide ranging exploration of multiple ethical issues in genetics, particularly human genetics. From cloning to embryonic stem cell research to genetic screening to designer babies, Polkinghorne surveys the landscape and even-handedly describes all sides of the argument. Interestingly, almost all the arguments he presents are secular, based on arguments like confidentiality, individual rights, children’s rights, and the greater good instead of the Bible or church tradition.

The section on stem cell research is somewhat of an exception to the secular rule. Here, Polkinghorne argues in favor of embryonic stem cell research by arguing that while medical procedures on a person should always be done in that person’s best interest, the embryo does not count as a full human person. A Catholic, he says, would take the dualist position that the soul enters the body at the moment of conception, thereby conferring the zygote full personhood right then and there. Polkinghorne’s psychosomatic view of the soul as an “information-bearing pattern” means that the soul develops continuously along with the rest of the body. In the early days of embryonic development, the embryo is an undifferentiated mass with no brain or nervous system—and hence no soul. The embryo is still entitled to respect because of its potential destiny as a full human person, but it is not yet a full human person. I find this psychosomatic view a lot more plausible than strict dualism, and more in accordance with lived reality.

The list of ethical issues Polkinghorne surveys is interesting in and of itself. Some, like the prenatal diagnosis of genetic diseases (and the consequent possibility of selective abortion), are just as relevant today as they were in 2005. Others are even more salient today than they were in 2005. Selective germline modification was science fiction in 2005, but in 2018, two CRISPR-modified baby girls were born thanks to the rogue scientist He Jiankui. Still other issues were hot topics in 2005, but are hardly discussed today. The chief example is human cloning.

What ever happened to human cloning? In my childhood, during the 2000s, human cloning was a hot topic. Everyone debated it. Governments around the world rushed to ban it. Clones popped up everywhere in popular culture, even in children’s shows. Now it’s 2020, and hardly anyone talks about human cloning. What gives?

[It turns out that although human cloning is possible with current technology, it is extremely difficult.](#) Dolly, the first mammal to be successfully cloned, was the only success out of 277 attempts. Recent advances in cloning efficiency have allowed

Chinese scientists to clone primates, and in 2018, the cloned crab-eating macques Zhong Zhong and Hua Hua were born. However, producing these two clones required 63 surrogate mothers and 417 eggs. This is still far too inefficient to make human cloning a practical possibility in the near future. This particular ethical issue, at least, is one for the future.

Imaginative Postscript: Some Naive Speculations

This chapter is my favorite. In this chapter, Polkinghorne uses charmingly “naive questions” (as he calls them) to explore the possibilities of the world to come, in much the same way that physicists use thought experiments to probe physical reality. This chapter is self-described speculation, which makes it all the more fun to read.

The first naive question is how the old and new creations—the material world and the world to come—are related. Polkinghorne conceives of them as parallel planes of existence, each with their own version of time. Sometimes, the two creations intersect and a portal to the spirit world opens in our world. This is what happened during the resurrection appearances of Christ. The risen Christ not only talked to his followers, but ate with them and allowed them to touch him, showing that exchange of matter between this world and the next is possible.

The second naive question is, what is the destiny of the matter in this world? Polkinghorne thinks the universe’s history will be allowed to play out. Since the universe seems destined for the eternal night of [heat death](#), when it becomes a uniform soup, God will probably wait until everything interesting stops happening before transmuting the matter into the matter of the new creation. All human beings, whose soul-patterns had been held in the divine memory for countless eons, will be given new bodies. We got a preview of this process of transmutation of matter and embodiment of souls with the resurrection of Jesus Christ, but the same destiny awaits us all.

The third naive question is, what will Jesus be like? Polkinghorne’s answer is unsurprising: Jesus will appear to us as authentically human, just like he did to his followers 2000 years ago. We will not be meeting Jesus one on one to shake hands, but all together as part of the corporate Christ; moreover, this meeting will be “communal and continuing”. In the same way that God is trinitarian in nature, the corporate Christ contains not just Jesus himself, but all believers, in a way reminiscent of particle-wave duality. The hive mind or the Borg might be good analogies if they did not have such negative connotations.

The final naive question is my favorite: what about the little green men? Surely if there is intelligent alien life out there, they deserve to share in the eschaton too? Polkinghorne thinks they will, because Jesus would have appeared to them as a little green man for salvific purposes. In the world to come,

Within the comprehensive community of the *totus Christus*, we shall encounter our Redeemer in the ways in which he chooses to make himself known to us. For humans it will doubtless be in the mode of authentic

humanity; for others it may be in the mode of authentic greenishness. For all, it will be in the fulness of his salvific bringing together of the divine and the created. Perhaps we may also hope to encounter each other in a corresponding openness and authenticity, so that humans and little green men will come to embrace and augment each other in the endless exploration of reality which is our ultimate destiny together.

What a beautiful vision, and how fitting for a scientist to have these eschatological hopes. I hope this vision will come true.

Conclusion

I highly recommend this book to anyone interested in the intersection between science and religion, whether Christian or non-believer. Polkinghorne is clear and concise, but not curt; he is firm in his Christian beliefs, but respectful of others. The book didn't change my opinion on Christianity, but it certainly increased my understanding of it. Considering the profound importance of Christianity to both world history and the world of today, understanding the beliefs of its adherents is essential to understanding the world around us.

Fantasyland

For future readers: this post was written just days after [pro-Trump rioters invaded the United States Capitol](#) to try to overturn the results of the 2020 presidential election).

I.

Amidst the social and political disintegration of the United States, my serendipitous choice of reading material is little consolation. That said, [Fantasyland: How America Went Haywire: A 500-Year History](#) by Kurt Andersen might have been the single best book to begin just a week before armed right-wing extremists invaded the U.S. Capitol.

Political violence, security failures, and even [democratic backsliding](#) are by no means peculiar to the U.S., but the rioters' delusional motivation made this extraordinary event a quintessentially American phenomenon.

Many protests - and yes, even some violent riots - are motivated by reasonable grievances arising from factual claims. Perhaps [Nat Turner's rebellion](#) was doomed to fail or counterproductive to the cause of abolition, but his and his followers were motivated by the very real cruelty of slavery (though in perfect Andersenian fashion, Turner "[was convinced that he 'was ordained for some great purpose in the hands of the Almighty'](#)"). Even ideologically-motivated violence like that of the USSR (or U.S. in Vietnam) often arises from an ethereal ideal like "equality" or "freedom" beyond the direct reach of factual verification.

II. A summary-summary

Fantasyland is an unapologetically sardonic account of America's legacy as a home for delusion and fantasy, not just in politics and religion but also, more interestingly, in aesthetics and leisure as well.

To avoid reinventing the wheel, here are some key quotes from [Wikipedia's book summary](#):

Part I: The Conjuring of America: 1517–1789

- *"Fantasyland contends that the earliest European settlers in what would become the United States were gold-crazed adventurers ([Jamestown](#)) and God-crazed cults ([Puritans](#) and [Pilgrims](#))"*

Part II: United States of Amazing: The 1800s

- *“The 19th century saw a mythologizing of the country's founding and founders, and a proliferation of religious sects”*
- *“Occult beliefs...were common”*
- *“The century introduced [homeopathy](#); medical fads, and [snake oil](#) peddlers*
- *“Steam-powered presses spawned large-circulation newspapers and magazines with loose standards of accuracy”*

Part III: A Long Arc Bending Toward Reason: 1900–1960

- *“Reason began fighting back: the [Pure Food and Drug Act](#) was passed... the [NAACP](#) was founded”*
- *“On the other side, there were brief conspiracy panics [and]... nostalgia for [Antebellum South](#) spread”*
- *“Fundamentalism in religion grew in popularity”*
- *“Movies...became more prevalent, bringing a greater amount of fiction into people's lives. There was an explosion in advertising and modern celebrity culture.”*

Part IV: Big Bang: The 1960s and '70s

- *“The 1960s and '70s were a time when bits of everyday life were being replaced with bits of everyday fiction, and there was a veritable explosion of [woo](#)-based ideologies taking hold”*
- *“[Woodstock](#), the [Counterculture](#), and [hippies](#) encouraged free thinking and finding one's own truth”*
- *“People could escape their mundane lives in living history theme parks, [Civil War reenactments](#), and [Renaissance fairs](#)”*
- *“There was a glut of theming: restaurants, malls, and architecture.”*

Part V: Fantasyland Scales: From the 1980s Through the Turn of the Century

- *“Reality TV was ubiquitous...Escapists could enjoy [Burning Man](#), [live action role-playing](#), [fantasy sports](#), and fantasy camps for adults”*
- *“The country had a [Hollywood president](#)...politics became entertainment”*
- *“The [FCC fairness doctrine](#) was eliminated, ushering in [Rush Limbaugh](#) and [Fox News](#)”*
- *“The [Internet](#) enabled every person access to every conceivable idea and interest”*

Part VI: The Problem with Fantasyland: From the 1980s to the Present and Beyond

- *“Fantasyland contends that two changes in American society led to modern tipping points: the [counterculture of the 1960s](#) and the [Information Age](#).”*
- *“The internet and world wide web permitted all manner of ideas to bypass the traditional gatekeepers of information dissemination.”*
- *“Conspiracy theories spread”*
- *“Fundamentalist Christianity is amplified in the GOP”*
- *“Extreme skepticism of the press is widespread as people lose immunity to false information.”*

III. The Good

1. The hidden connection

Andersen’s most basic insight is his unification of disparate strands of American history and society. What unites the Salem witch trials, Disneyworld, WWE, Kuhn’s [The Structure of Scientific Revolutions](#), and the NRA? Fantasy (or perhaps more fittingly, antirealism) - for better and for worse.

This unification is particularly insightful when the elements stand in apparent opposition. The relativism of 1980s academia may have been culturally, intellectually, and ideologically distant from both left new-age spiritualism and right Christian fundamentalism of the same era, but all three trends mutually reinforced a cultural acceptance of *belief* as a legitimate metric of truth.

2. Reframing culture

I also appreciated Andersen’s reframing of (apparently) benign elements of American aesthetics and culture. He describes the long history of American nostalgia for some idealized past - in other words, a desire for something that does not exist and may have never existed.

For example, suburbs are a product of our simultaneous desire for modern urban amenities and a “pastoral” aesthetic and lifestyle, and from the 1980s “new skyscrapers and shopping malls and residential buildings were practically required to reproduce... styles from other eras or continents.” Likewise, Americans’ love affair with firearms and SUVs is a means of [LARPing](#) some image of the rugged outdoorsman.

An Aside: Boy Scouts

This framing shed real light on part of my childhood experience as a Boy Scout that I’ve had trouble putting into words. At an obvious level, the camping trips and outdoors/survival skills reflect a cultural appreciation of, even nostalgia for, this “rugged

outdoorsman” ideal. I was already well-aware of this; if some boys like camping and building fires, the subtle historical-cultural factors at play are interesting but mostly irrelevant.

More subtly, Boy Scouts has an uncanny and jarring relationship with anachronism. Perhaps a few examples will best convey what I mean:

- The guy who keeps track of equipment is called the “quartermaster.”
- The advancement ceremony is called a “Court of Honor”
- Everything - buildings, uniforms, signs - looks *intentionally* old. Not from a lack of ability to be replaced, but as an intentional decision to foster an image of a bygone era.



I spent many nights in one of these

This aesthetic is most obvious in settings of centralized planning and organization, like the summer camps and [national bases](#).

The point is this: the Scouts’ style and curriculum are not merely old-seeming as a *byproduct* of the functional intentions of the program (to teach life lessons, socialize boys, have fun, etc.). They represent a non-functional effort to recreate some supposed historical ideal, just like so many elements of American culture. A few analogous examples:

- The design and marketing surrounding Jeep SUVs and similar cars gesture towards a lifestyle of safaris and outdoor exploration that 99% of their owners will

never use their cars to partake in.



- Quasi-functional shoes like Vans (skateboarding) and Sperrys (boating, whatever that means) likely at one point signalled their owners' genuine participation in these activities, and have since transitioned to purely nonfunctional form reminiscent of a bygone era.
- Camden Yards, home of the Baltimore Orioles, is intentionally designed to look and feel old.



None of this is to disparage anyone who buys any of these items; I own Vans and have gone to Orioles games. At a larger level, though, it may indicate cultural preoccupation with nostalgia.

3. Combining real and fake

Fantasyland's third interesting theme is the frequent combination of fantasy with reality or credibility, mostly as a means of lending legitimacy to the former. Conspiracy theorists go out of their way to find and present genuine evidence alongside complete fabrication. Climate deniers find (or finance) people with PhDs to go on CNN and Fox. WWE stars incorporate unplanned events in their real lives into matches' reality-show storylines.

While such combinations need not be deceptive or harmful (as is often the case in entertainment), untruths can be much more nefarious and powerful when bestowed with undeserved credibility. There is a reason why the anti-vaccination movement massively accelerated after a paper linking vaccines to autism was published in a prestigious journal (and [has since been retracted](#)).

4. Americans believe crazy things

After point 1 above, my second biggest takeaway is that Americans believe some kinda crazy shit. I know that some people [might answer polls in bad faith](#), and that recent political polling errors have cast doubt on the validity of other forms of polling that never have their answers checked in real life, but even still. A few examples:

- *“one in nine adults, 25 or 30 million Americans, are sure they’ve ‘experienced/witnessed the devil/evil spirits being driven out of a person’”*
- *“in the fall of 2016...an Economist/YouGov survey found a majority of Republicans still believed Obama probably or definitely was born in Kenya.”*
- *“Only a third of us...believe with some certainty that CO2 emissions... are the main cause of Earth’s warming...[and] more than a third of us believe that...it’s a hoax perpetrated by a conspiracy of scientists, government, and journalists.”*
- *“Only a third strongly disbelieve in telepathy and ghosts.”*
- *“A quarter believe that vaccines cause autism and that Donald Trump won the popular vote in 2016...”*

These proportions remain pretty high even after we subtract Scott Alexander’s proposed [4% “lizardman constant.”](#) Other beliefs or practices aren’t so much crazy as merely surprising, at least for someone in a liberal, educated bubble like myself:

- *“A majority of Americans tell Pew they pray every day; in the rest of the developed world, those fractions are one-tenth or one fifth.”*
- *“In early 1998, as soon as we learned that Clinton had been fellated by an intern around the Oval Office, his popularity spiked, according to the polls.”*

And perhaps most importantly:

“When I say a third believe X or a quarter believe Y, it’s important to understand that those are different thirds and quarters of the U.S. population.”

It’s fashionable and important to decry naive [scientism](#) and identify the hidden functions of folk wisdom (as I discussed in a [previous post](#)), but it can be easy to forget that a lot of American adults believe things far outside the intelligentsia’s Overton window.

(A thought: I can imagine some of my own beliefs/things I consider plausible being characterized as “crazy shit” by someone writing a different version of this book: panpsychism, the simulation hypothesis and concern about human extinction from AI, for example).

IV. The Bad

Some people might be put off by the author’s contemptuous tone, which I suspect is a product of [Andersen’s history](#) as a novelist and humorist. In fact, much like in [Nassim Taleb’s](#) work, this may be part and parcel of his ability to clearly and unapologetically state his case.

In comparison to his mocking of socially-harmful practices or beliefs, I’m much less sympathetic to the antipathy aimed at genuinely benign forms of leisure. Who cares that “at least a third of the people at theme parks are adults without children” or that “Walmart alone sold \$200 million worth of coloring books for adults in 2016” or that folks with families and jobs go to Comic-Con or LARPs or Civil War reenactments? Life is hard enough. *Let people enjoy things.*

Ultimately, though, this is a matter of taste and doesn’t bear directly on any of his arguments. Here are a few more substantive disagreements or issues I have with the book:

1. Lack of Explanation

While the book’s mass of historical evidence and comparison between the U.S. and other countries makes a convincing case that fantasy/deception/pretend/subjectivism has been a persistent and important element of American culture, surprisingly little ink is spilled on explaining the more fundamental reasons *why* this is the case.

If I had to articulate an explanation espoused by the book, it would be that this ethos was instantiated at America's founding by its religious fundamentalist and gold-seeking settlers, and has proliferated due to something like memetic or cultural evolution.

However, this is never stated explicitly, and other non-cultural explanations are hardly considered. For instance, Andersen at most vaguely gestures toward the theory that the U.S. has absorbed immigrants particularly genetically-predisposed to believe in myth, and thus has descendent citizens with similar inclinations. While I have no idea if this theory holds any water, there is a whole subfield of "[dual inheritance theory](#)" studying the interaction between genetic and cultural evolution.

Likewise, I can imagine many plausible economic, geographic, or political factors at play; perhaps America's large amount of unsettled (by white people) land for much of its history was partially causally responsible. Perhaps its relative economic and political freedom and/or economic mobility (again, for some of its history and for some citizens only) was somehow causally responsible.

Maybe Andersen thinks that some or all of these things are obvious and implicit, but that's not the sense I got from the book.

2. Out-of-hand dismissals and misrepresentations

At several points, Andersen seems to reject some concept or proposition as incorrect without due consideration or misrepresents the claims of some text, scholar, or subfield.

In particular, he dedicates most of an entire chapter to academia's new espousal of "relativism" beginning in the 1960s, and how "postmodern intellectuals...turned out to be useful idiots for the American right." While the symbiotic relationship between various forms of relativism and rightism is interesting and plausible, Andersen fails to even consider whether some of these intellectual "Squishies" might actually be correct.

Don't get me wrong; I think that many propositions are definitively true or false, and even tentatively believe in some form of [moral realism](#). However, Andersen seems to conflate some very different forms of [epistemic relativism](#) - namely, one in which a single proposition is in fact true for one observer/culture/society and false to a different one, and another (much more plausible) version in which there are no observer-independent means of *ascertaining* whether a statement is true or false. For a good intro to the latter position, see [this 1982 paper](#).

It is completely possible both that the second form of relativism is correct and that it has had a negative impact on American society, but that's not what you'd get from reading *Fantasyland*.

3. Excessive credulity

Perhaps my biggest issue with the book is its relatively uncritical stance towards science and expertise. In framing many harmful trends (e.g. anti-vaccination, creationist

education) primarily as antithetical to institutional authority, Andersen consistently implies that experts and scientists are the “good guys” living in “Realitystan” adjacent to Fantasyland.

In many circumstances, this is basically correct; mainstream medicine, for example, is almost certainly correct that vaccines are safe and good and effective. Other times, not so much. Large segments of social science are reckoning with the realization that their fields’ empirical foundations are misleading or incorrect (and if you don’t know much about the replication crisis, I highly recommend [this piece](#)).

As has become salient recently, public health experts basically [lied](#) (or if you prefer, misrepresented the quality of evidence) about the effectiveness of masks in preventing COVID transmission, siphoning away much of their hard-earned credibility.



U.S. Surgeon General  · 2/29/20 

Seriously people- STOP BUYING MASKS!

They are NOT effective in preventing general public from catching [#Coronavirus](#), but if healthcare providers can’t get them to care for sick patients, it puts them and our communities at risk!
bit.ly/37Ay6Cm

 6,556  42.9K  67.9K  

I also think it’s safe to say that nutrition as a field is largely in the same boat as social psychology. As Alvaro de Menard writes in “[Are Experts Real?](#)”

“nutrition is pretty much completely fake, yet the government dutifully informs you that you should eat tons of cereal and a couple loaves of bread every day. A USDA bureaucrat can hardly override the Scientists...”

In fact, there’s a good case to be made that “[Most Published Research Findings Are False](#)” (a paper with its own [Wikipedia page](#)!)

In ignoring how institutional authority - the closest thing to the book’s protagonist - *itself* often lives in Fantasyland, Andersen presents a biased portrait of America’s epistemological landscape.

V. Conclusion

I'd recommend *Fantasyland* to anyone willing to approach it with a humble and skeptical stance, with bonus points for reading it alongside thoughtful objections to naive rationalism/empiricism/scientism (e.g. "[Are Experts Real?](#)" "[Relativism, Rationalism, and the Sociology of Knowledge](#)," and [The Structure of Scientific Revolutions](#) linked above; [The Black Swan](#) and [A Secular Age](#)).

To reiterate a point from section III, it can be easy to get caught up in these sorts of interesting and thoughtful objections. *Fantasyland* serves as a reminder that folk wisdom *is* sometimes just wrong and nonfunctional, experts and scientists often *are* more correct than the masses, and many of our fellow Americans believe in things like [QAnon](#), exorcism, the causal power of prayer, and vaccine-induced-autism.

Having never enjoyed studying history in its own right (though I wish I did), this book also made for an entertaining reintroduction to many of American history's most important social and political currents. If you want to understand the legacy behind this week's insurrection, *Fantasyland* is the place to go.

Farmers for Forty Centuries by F. H. King

I first read this book that described an early 1900s tour of predominantly pre-industrial agriculture in Japan and China over a decade ago and it made a very positive impression on me at the time as a budding experimental farmer. Revisiting the work has cast it in a different light in my mind, though I will summarise and analyse the most interesting parts of the book.

The work outlines a journey through a landscape thoroughly dominated by human activity. In the wetter south of China hand dug canals crisscrossed the landscape, allowing irrigated rice and other crops to be grown intensively throughout the year. In the drier north of China this activity was restricted to river flats, with the drier fields hand irrigated from hand dug wells during droughts. A family was observed here planting sweet potato shoots in dry soil while the father carried water in buckets from a ravine 500 m away. Any steeper land was continually cut for firewood and mulching material, which was carted back to the cities and fields below, leaving a landscape mostly denuded of trees. The flat landscape was dotted with man-made hills created to house the graves of ancestors.

Frequent discussions were made with the density of humans upon the landscape, estimated at 2 acres per person (with only one acre of that arable) compared to 20 acres per person in the USA at the time. Japan at this time already had 0.3 arable acres per person. Today this has since shrunk to 0.2 acres of arable land per person in China, and 1.2 acres arable land per person in the USA for comparison (close to the global average). The author estimated Japan had 646 times the density of all animals and humans than the USA at the time (likely supported by the high dependency on seafood in Japan). These extraordinary population densities were supported by rice, with wet cultivation in paddies producing double the yields per crop and more crops per year than unirrigated upland rice. The average Japanese family farmed 2.5 acres of irrigated rice, with small paddies often scattered around the village. Foot pumps were used to raise water from the canals to flood the paddies. Other crops were grown on the paddy walls.

By starting rice seedlings in nurse beds the time the crop spends in the field is reduced, reducing irrigation and weeding labour (with the flooding acting as an effective weed suppressant), and ensuring the young seedlings rapidly absorbed the fertiliser immediately after transplantation. Legumes were grown in the fields before flooding, providing nitrogen, and transplanting gives them several extra weeks to grow. One person can transplant a third of an acre in a day, yielding about 300 kg of rice. Just taking into account the transplantation labour on this space the EROEI (energy return on energy invested) is 191. If you include similar amounts of time for field preparation, irrigation, harvesting and processing the EROEI drops to 24, still enough to support a fairly complex civilisation. Annual rice consumption was about 136 kg per person, providing about 24% of total energy needs.

Paddies rented for 400 kg rice per acre, and a crop typically brought in 400 kg more for the farmer. With rice giving 5600 kJ per kilogram, this represents 2 240 000 kJ or 257 days' worth of food. Most families farmed about 3 acres of rice, giving 770 days' worth of rice energy per year,

enough to feed 2 people eating nothing but rice, or 8 people with rice providing one quarter their calories as was more typical in the region. Families owning their own land would get double the benefit from rice farming since they wouldn't be paying half the crop as rent, so 16 people could be reasonably supplied rice by 3 acres of paddy, doubled again in places where two crops per year are possible.

These cultures also relied on consumption of large quantities of vegetables. Pigs and chickens were more common than cows and sheep since their feed conversion efficiency is much higher. Large flocks of ducks and geese kept in the wet lowlands on boats in the canals, herded out at each day and back to the boat at night allowing easy movement to new grounds. In the drier north herds of goats and sheep graze the roadsides and grave land hills, refusing to take a single bite of unfenced crops growing in the fields. Only boiled water was drunk due to disease risk from high population densities (one reason for importance of tea culture). I wonder what the effect of constant caffeination was on peasant productivity, since a comparable cultural transformation in productivity took place in Europe with the introduction of tea, chocolate and coffee at the start of the industrial revolution.

Nutrient cycling was practiced to a fanatical extent. All human and animal waste was collected in concrete lined pits and mixed with other vegetation. Almost zero flies were observed during the entire visit. Once fibre was broken down it was dried, mixed with ash then applied to the fields. Japan averaged 1.75 tons of this compost per acre of crop land, all processed and transported by hand. Liquid manure was also often carried out in buckets and applied with a long handled dipper. Canal mud was also frequently reclaimed and applied to fields at around 70 tons per acre, equivalent to a 1 cm deep layer. Flooding in the lowlands periodically made the canals impassable, preventing compost transport. Droughts and deforestation in the north caused massive erosion of denuded hillsides, though the author observed them being protected and recovering in places (presumably due to the growing use of coal). Firewood seems to be the most limiting resource for this society with people being extremely frugal in its use. Modern societies have been ignoring forests near human habitation for a century with massive regrowth. I wonder how rapidly it would be all cut down again if we were ever forced to return to using fire wood.

Some interesting technologies were observed, including the foot powered water pumps used for irrigation that consumed many hours of labour. Freight wheel barrows with a single large wheel and the load balanced either side of it were used to transport goods and people long distances with just a single person pushing up to a dozen passengers. Where possible animals and sails were used to add more power. The generally flat landscape made these practical. Human porters were often used for smaller loads in more difficult situations.

The picture painted of this society might sound like the ideal end point for permaculture, or at the very least an important source of inspiration. Closer reflection on the realities of this world might cast things in a different light. The author noted that mothers typically had 7 to 11 children, but only 1 to 3 lived to adulthood. Apart from the usual scourges of epidemic diseases made possible by such high population densities infanticide was commonplace, especially for

girls and especially during famine years. Famines in China were recorded at an average rate of once every year for the last two thousand years, with either drought in the north or flooding in the south. Major wars and battles also occurred every few years. The period of scientific breakthroughs from 618-1279 gave way to a prolonged period of stagnation, eventually allowing invasion by Europeans. The end of this golden age is when wet rice cultivation was developed, causing a massive population boom in the previously underdeveloped south. The massive food surpluses created population surpluses, making labour very cheap, and creating a disincentive for the further development of science and technology. The author noted with astonishment how all sorts of labour in China could be bought for a tiny fraction of the price in the USA. If permaculture achieves its central dream of creating highly productive and sustainable agricultural systems what is to stop the same culture of never ending toil and exploitation befalling the people?

The take home message for me isn't the state that the author experienced pre-industrial agriculture, as wondrous and terrible as it was. Instead it makes me cast my mind back to the first peasant farmer who shovelled some mud around and created the first paddy, then another who discovered that rice tolerates transplantation and flooding, then another who discovered the legumes that can grow in the fields before flooding, then the tinkerer who developed the foot pumps. None of these people could know the immense power of their inventions when combined together, nor the unforeseen consequences, both good and bad. During our current golden age of industrial plenty and leisure we have a rare opportunity to do similar experiments, confident in the fact that we won't starve if they don't work out to our immediate advantage. For me permaculture is about using this brief period in history to do the difficult experiments that nobody driven by hunger or profit is going to make, and seeing where they take us in the thousands of years to come.

Feed by M.T. Anderson

M.T. Anderson admits *Feed* is more contemporary allegory than forecast for the future.¹ The novel succeeds in illustrating the threats mass media, technology, and Big Corp pose to independent thought and the longevity of the human species and while it may not thoroughly discuss underlying causes, it certainly does nail the symptoms: the destruction of the family unit, the extension of childhood into stillborn adulthood, the pervasiveness of ironic language, the breeding of people as consumers through schooling and media, the ecologically intertwined world war, and the mechanical replacement and subsequent atrophy of organic limbs. Published in 2002, before Twitter, TikTok, and YouTube, its concerns have, unfortunately, only grown more relevant today and with Elon Musk pursuing Neuralink, the "feed" itself is not even allegory anymore but literal near future reality.

The troubling confusion with *Feed*, however, is that Anderson shares these concerns by participating as celebrity author in the commodification of young minds with a coming-of-age book of his own, which, it should be noted, fails to offer any resolutions other than death or depression. Imagine the pitch to Candlewick Press: "It's like *Catcher in the Rye* but Sci-Fi!" The intended disappointment, when Titus concedes trying to break through the alienation of consumerism, is, of course, frustrating, but furthermore, on the conceptual plane, Anderson's dystopia is particularly disgusting because it is Zion-less. Violet, the character with the greatest sense of agency, appears to comprehend the horrors of techno-capitalism and *Feed* reads as if its audience should admire her materialistic self-control. Yet her stoicism is ultimately empty, as she never dares to pursue anything more sublime than "going up to the mountains" with Titus. Violet's father appears to offer tools for independent thought. Yet rather than provide love and guidance, this potential mentor only scorns (*You don't even get my Time Machine reference! Pathetic child!*). Is this Anderson's attitude toward the next generation, too? Rather than a dynamic journey from ignorance to inspired purpose, *Feed's* endorsed heroic rite of passage is stagnant lolling in shame, self-pity, and hopelessness. Titus may not solve his problems, but, look, he has officially matured, because now he reads the worldly news like Violet and is sad about it, too! Even worse, Anderson provides little reason to even seek revolution. Violet's hysteria regarding the imminent end of the world goes unopposed. So, too, does Titus' anti-climatic realization that "there's nothing more [to life] than that... just us, on the bridge of dreams." Why bother fighting for change, then? There is an argument for realism, that these forces are incredibly complex and that it does more harm than good to ignore, sugarcoat, or simplify. But then again, dreams of "bubble neighborhoods" and "upcars" are tolerable, but a character finding true independence and meaning outside of mainstream America is too implausible? A character having hope for the future is off limits?² The real danger *Feed* poses is that its audience will not grasp the a priori implications of the book's very existence: that it was written by an actual American adult, accepted by a major publishing company, printed thousands of times, distributed to stores and schools across the country, and brandished, even, with a shiny silver emblem. The deeper message kids *should* take away from *Feed* is that their parents are spiritually broke.

But what if Anderson is aware of this irony? Like saying Coke™ one thousand times is funny *if you know it's an advertisement*. As in, Anderson demonstrates just how perversely doomed America is by writing *an award winning* young adult novel that only affirms modern life as meaningless and destructive. How should we judge this National Book Award Finalist then? One way would be by its material effect on the world, regardless of intention.³ It is still unclear

what good *Feed's* publication could have spawned, even if its supposed satiric end were easily intelligible and not merely attainable by the kind of hypothetical leap taken here. Should an uprising of Freshman Literature teachers bearing red pens have stormed the offices of those superintendents who forced this gross material upon their classrooms? It should not have been lost on *Feed's* author that an ironic fight against mainstream America could only achieve that of Violet's own futile protests. On the other hand, Musk and the like actually provide a value system as intriguing as anything else available in the mainstream media market. Despite the antics, Musk's pursuits are founded in the beliefs that the human species is dignified and precious and that our most difficult problems should be attacked with bold aspirations. The philosophies embodied by Silicon Valley are not flawless, but the road to *Feed's* hell *is* paved by compelling visions. Even more reason for those who fear the externalities of these pursuits of infinite profits and the perfectibility of man to provide competitive alternatives. Anderson, however, appears too concerned with the lexically appropriate and Netflix-able to offer any. Either that or he has, himself, given up. Why write young adult fiction, then?

[1] He says so in the interview appended to my edition of the novel (ISBN 9780763622596).

[2] If we are to find solutions, don't we have to first believe they exist? [Here's](#) physicist David Deutsch on optimism.

[3] Philosophy Tube has [a nice breakdown](#) of Slavoj Žižek's ideas on this approach.

Fitzpatrick's War by Theodore Judson

A lot of people enjoy thinking about - fantasizing about, if we're being honest with ourselves- the apocalypse. For whatever reason, the concept of the current world order crashing and burning and being replaced with anarchy and violence is a pleasant thought for a lot of audiences. There is undoubtedly plenty of good chunks of Psychological Significance Ore to be mined from figuring out why this might be, but one aspect I rarely see in modern depictions of the world ending is that the world has already ended many, many times and the human race kept going anyway, and those crises resulted in new civilizations with new cultural modes that seemed radical and alien to those that came before, even as the next few hundred years shall very likely seem horrifically bizarre to us living today.

That's the problem with history; it doesn't end. It just keeps pouring out from the primordial slime of the distant past until it ends up at your doorstep, and then it crashes through your walls and out the back while you're still trying to get a handle on things.

Fitzpatrick's War is a 2004 novel by Theodore Judson. In it, the modern, globalist world ended in a series of disasters- plagues, nukes, EMP terror strikes, civil war- and the rightwing tradcon survivalists inherit the earth. The plot of the novel, as compelling and relevant as it is, is almost secondary to exploring the society that these militant rednecks created for themselves, and how this society responds to the world outside of its borders.

The world of *Fitzpatrick's War* strongly resembles a nightmare from the alt-right psyche (or a daydream, depending on how accelerationist the dreamer is). The world is divvied up into 3 major powers- communist China which dominates most of East Asia and the Pacific Islands, a Pan-Islamic Caliphate that rules the Middle East and most of Europe, and the Yukon Confederacy, which has forged a new version of Western civilization that incorporates the continental United States, Australia, Canada, and the British Isles (in this universe Brexit clearly did its job and did it well). There are also some minor powers- several bandito kingdoms and mini-empires in South and Central America, Pan-Slavia (a rump state in what was once Russia, where the old Soviet bloc still resists the Muslim horde), and India. Sub-Saharan Africa gets screwed with nonstop civil wars as each superpower uses it as a proxy battleground- the Caliphate arms any group calling itself Muslim, the Yukons arm any group calling itself Christian, and the Chinese arm any group spouting Marxism.

Since the vast majority of this review and most of the book itself revolve around the culture and structure of the Yukon Confederacy, it is worth investing some time in describing it in some detail. The Yukons started as a loosely organized network of survivalists, Mennonites, farmers, and blue collar engineers spread throughout the world. In the course of the long and drawn out apocalypse, the Yukons coalesced into a unified front and violently purged their countries of the Mad Max style raider gangs and

the last remnants of the legitimate government. To end the civil war, the Yukons' pet engineers unleashed self-perpetuating EMP storms to permanently knock out all technology more complicated than a steam engine across the entire globe, thus giving the rural and mechanically inclined rednecks a dominant edge forever. The day they marched on Washington to give the treacherous, immoral, arrogant lawyers in Congress the bayonet is their Independence Day.

The Yukons are, as a people, deeply reactionary. They hold the family to be sacred, dividing themselves up into clans based on blood and marriage. They have a rigidly stratified society based on some manner of neomanoralism, with Lords and Barons and Dukes owning land and permanently renting it out to small farmers within their domain (this neomanorialism is in practice less oppressive than it sounds, since almost every renter is a blood relative of their landlord- I'm given to understand that historical manorialism was often like that as well). They are extremely militaristic, with every male not needed on the farm signing up for the Army and beginning their combat training before puberty even begins. They are also deeply devout Christians, although they tend more towards an Old Testament "smite the heathen hip and thigh" style of worship in contrast to the "Sermon of the Mount" flavor. The Yukon women are diligent mothers-and-wives-in-training until inevitable marriage at age 25 (no older, no younger, for inheritance reasons), at which point they become diligent wives and mothers and start raising the next generation of Yukon killers and breeders. Like the British Empire whom they consciously mimic, they fetishize Greek and Roman philosophy and history to the point where mastering Cicero and memorizing the *Anabasis* in the original Greek is considered a major component of an officer and a gentleman's education. They are also super, super racist- their society is explicitly and unapologetically white supremacist, to the point where even hinting that a man might have some Jewish heritage is grounds for a duel to the death. Every racial minority is relegated to the fringes of society doing work that is considered dishonorable for real Yukons, and even when the coloreds know their place, the police still hassle them.

All of this sort of blends together into an *ethos* that has been transmitted from one generation to the next for hundreds of years. If you are familiar with *Albion's Seed*, imagine a culture wherein the fervent rectitude of the Puritans was cross-pollinated with the savage xenophobia of the Borderers. To wit: the Yukon Confederacy is special. We are chosen by God to represent his light in the world. We cannot falter, we cannot become corrupt, and we can never, ever lose a war or else the scary foreigners will wipe us off the face of the planet and God's light will be extinguished from the world. Other empires have risen and fallen in accordance with Man's sinful nature; that fate must never befall us. As long as we keep ourselves pure and show courage in adversary, God will always deliver us from the evil foreigners. So praise the Lord and pass the ammunition.

By the time that *Fitzpatrick's War* begins in the early 2400's, the Yukons have settled into a comfortable rut. They leverage their superior technology to produce most of the planet's medicine, food, and tools, and enforce their economic monopoly by serially waging wars against the Yellow Menace of China, the heathen Muslims, the Mexican

banditos, and anyone else who looks at them funny. Their only allies are India and the Pan-Slavic remnant. Since the Timermen (the same engineers who had once ruined the sky and delivered the world into Yukon hands) control the only remaining satellites, and these satellites allow for precision air strikes and advance warning of enemy movement, the Yukons win every war they fight. However, they do not fundamentally unbalance the world order doing it- think of it like hunting conservation, almost. They never kill people at unsustainable levels, and that means their sons and grandsons get to go to war with the same enemies their dads and granddads did.

From this pattern emerges the plot of *Fitzpatrick's War*.

The book is structured as the memoirs of an elderly and thoughtful Robert Bruce, a Knight and hero who had served in an epic world war in his youth. He was the close friend and servant of the eponymous Isaac Fitzpatrick, the greatest statesman in Yukon history, who had set out to conquer the entire world and actually succeeded within about half a year. Every schoolchild knows and venerates the beloved and assassinated Fitzpatrick, but Bruce was actually there and saw the dark side of the hero that never made it into the history books.

Judson skillfully derives some humor from the format. Bruce's memoirs are presented and fact-checked by a bigoted and vain academic professor who keeps butting in with footnotes, insisting that certain events (such as war crimes, corruption, and religious deviance) that Bruce describes were either misinterpreted or are baseless slander against the great Fitzpatrick. Also, as Bruce begins questioning some of the tenets of Yukon life, all of the ethical choices that Bruce makes in accordance with *our* morality-mercy instead of brutality, love instead of ambition, egalitarianism instead of bigotry- are used as evidence of Bruce's untrustworthy nature. Any man who would refuse to execute helpless prisoners of war would obviously lie about a saint like Fitzpatrick; a man who would dance with his wife like a sissy boy is too morally bankrupt to believe when he makes accusations of patricide.

Games People Play: The Psychology of Human Relationships by Eric Berne

Have you ever had your boss pass their mistakes on to you?

Well, Bob the Boss does. At a critical meeting, Bob asks his team for suggestions on how to fix things. Alice shares an interesting idea, and Bob takes it to upper management. However, it ends up making things worse. Tough luck. Bob then redirects all the blowback to Alice. “Look what you made me do, Alice! Fix it, it’s on you now”.

Bob is playing a game of “Look What You Made Me Do”. He’s setting things up so the blame never lands on him. He’s vindicating himself. If you can relate to this game, you’re not alone. There’s hundreds of games people play, and psychologists have spent years studying these games.

However, all is not lost. If you figure out the game and play the right moves, you can stop these games.

- You can stop your manager from shirking off responsibility.
- You can stop your spouse from playing status games.
- You can stop your friend from messing up their life and seeking sympathy again and again.
- You can stop strangers from playing games with you, too.

But that’s only if you know the game they’re playing.

[Games People Play](#) is a repository of such games and how to stop them. Eric Berne spent 40 years studying these games and published an academic book. I spent a month plowing through the literature, took the best parts and stitched together a narrative of the most important games, with examples relevant to today’s age.

We'll start with some background theory — just enough to make sense of things. Then, we'll explore interesting games people play, like the one above, and figure out ways to stop them. If you don't want the background, you can [skip directly to the games](#).

Background

To understand games, we need to understand how people behave, and how this behaviour changes in front of others.

People have a set of behaviour patterns that correspond to one state of mind, and others that correspond to another state of mind, which are often inconsistent with the first. These different states of mind are ego states.

Eric Berne, the author of *Games People Play*, says there are 3 big ones: The Parent, The Adult, and The Child. These ego states are different modes of operation in every person.

The Adult is the rational logical self. The Parent is the caring, taking care of someone self. The Child is the playful, creative, easily offended self.

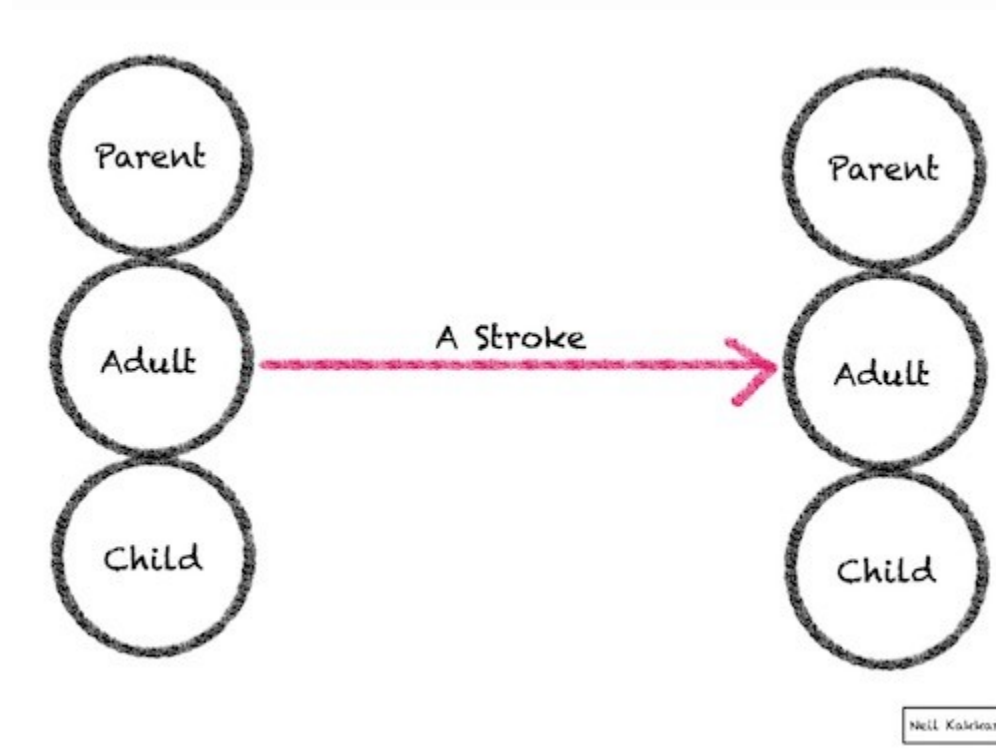
Note that these states have nothing to do with being a child or a parent yourself. A child acts from the Parent state when she tells their parents "I didn't like it when you did this. Please don't do it".

Sometimes, you can notice people transition from one state to another in conversation. For example, when Alice and Bob are discussing a problem and Bob claims that Alice is wrong, then Alice might get triggered out of the Adult state to a Child state where they start crying, or create an uproar, or call their mother to tell how Bob is wrong and mean. The transitions show up in physical gestures and posture too.

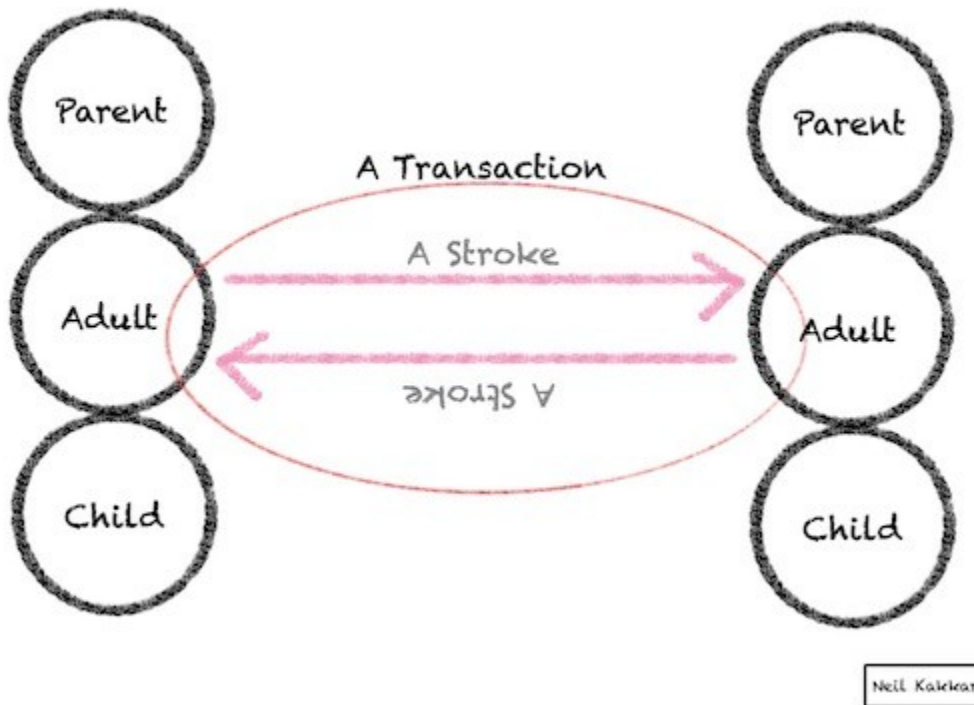
To understand people, you need to understand their ego states. To understand ego states, you need to interact. This is the core idea of [Transactional Analysis](#). *Games People Play* are built on top of Transactional Analysis.

Transactional Analysis

A stroke is a basic unit of conversational stimulus. Saying "Hi!" to someone is a stroke. Every stroke goes from one ego state to another.



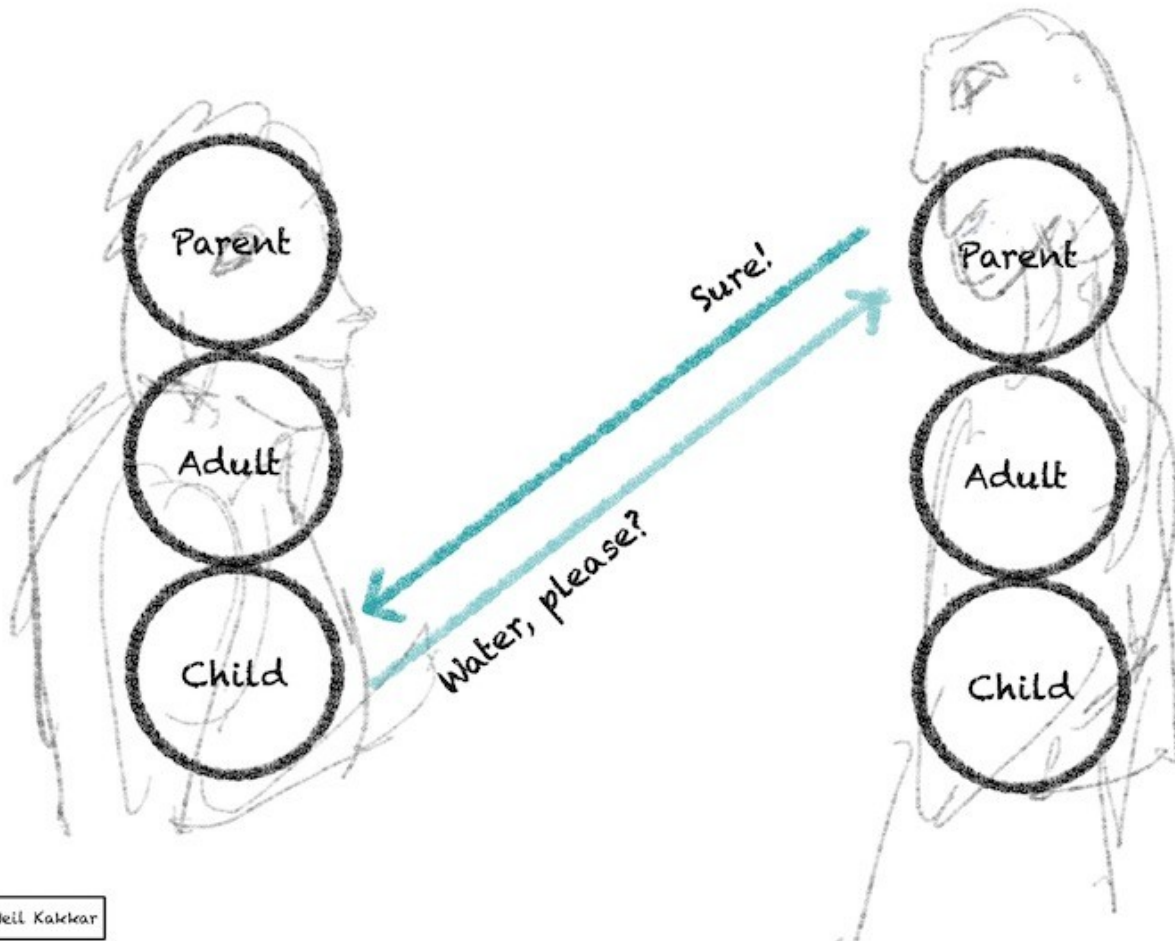
Multiple strokes together make up a transaction: a conversation between people.



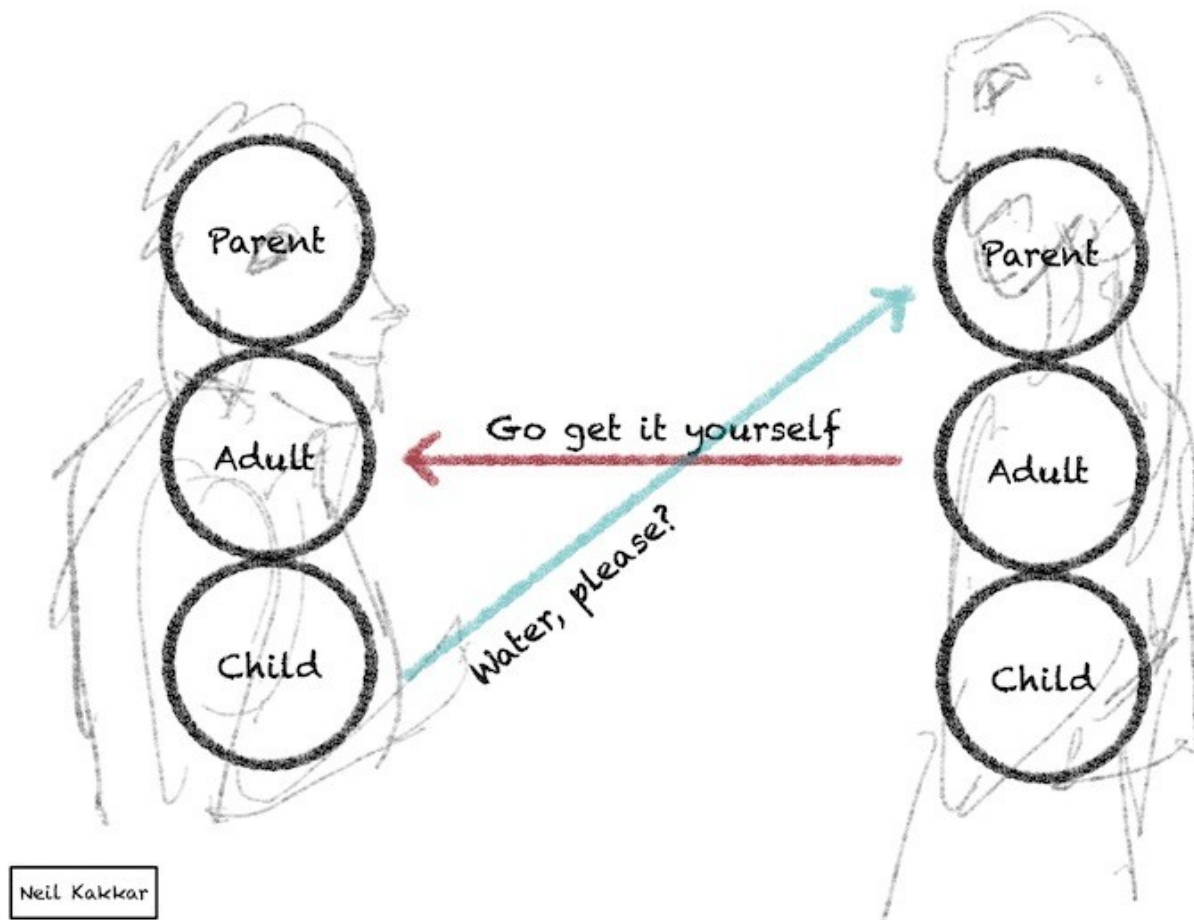
Like we need food to stay physically healthy, Berne claims we need strokes to stay emotionally healthy. For example, a movie actor may require hundreds of strokes each week from anonymous and undifferentiated admirers to keep himself going, while a scientist may keep physically and mentally healthy on one stroke a year from a respected master.

Now things get interesting. Every stroke expects a complementary stroke. Communication can proceed indefinitely, as long as the strokes are complementary: Adult-Adult, or Parent-Child & Child-Parent.

For example, consider a fevered child asking for water.



However, if you're sending a stroke to someone's Child state, and their Adult responds, conversation would break. This is a crossed transaction.



That's all we need to dive into games. There's more in the book, but I've covered all the main points.

Games

When any social interaction has an ulterior motive, and a payoff, it is most likely a game.

Berne writes: "If someone frankly asks for reassurance and gets it, that is an operation. If someone asks for reassurance, and after it is given turns it in some way to the disadvantage of the giver, that is a game."

The ulterior motives of Games are hidden transactions. On the surface they seem like normal interactions, but psychologically, they're different. For example:

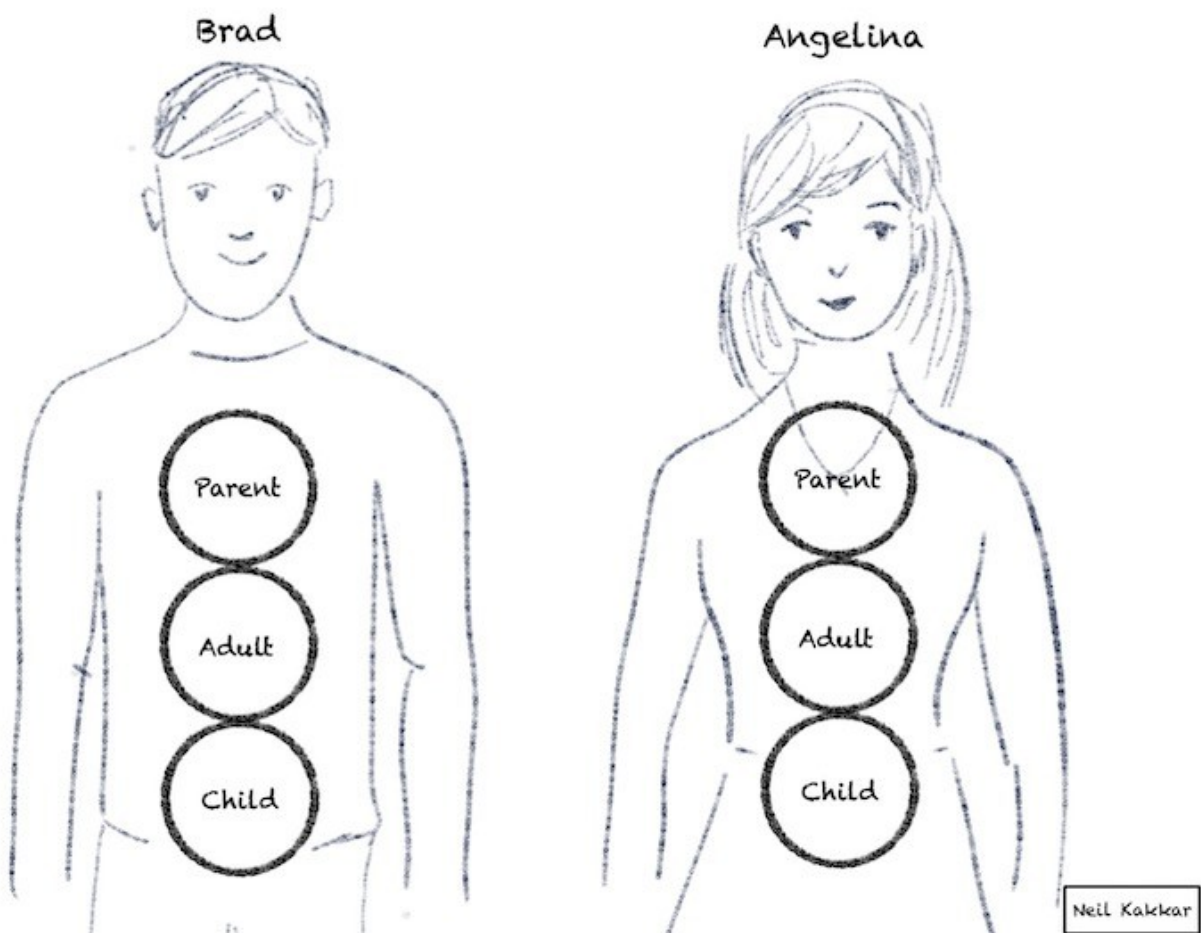
“Salesman: ‘This one is better, but you can’t afford it.’”

“Housewife: ‘That’s the one I’ll take.’”

While on the surface it’s an Adult-Adult transaction, psychologically, the salesman is instigating the Child of the housewife.

Now, for a few games. Meet Brad and Angelina, they’ll be our players throughout this blogpost.

Game - Examples



I'm only trying to help

Brad defers a lot of decision-making to Angelina. He tells her that you should help make the right decisions for our kids - you know what's best, and I'll help implement those decisions. One day, Angelina decides to not let the kids eat candy and coke anymore. So, Brad takes away their candies and coke. This upsets the children, which makes Brad upset. He tells Angelina, "Look what you made me do, now the kids are upset!". This bewilders Angelina, and she cries out "I was only trying to help!".

The game is based on the supposition that Brad is ungrateful and disappointing. The payoff comes with the bewilderment: when, despite all your help, he is ungrateful.

It's a form of a bigger class of games, where you're punishing yourself for something that might've happened, and reaffirming your position that the world is ungrateful, and you're a saint for still trying to help people out.

The game isn't just played by couples. It shows up at work, too, when a colleague gives advice on what to do in a difficult situation, knowing that the other person wouldn't follow it.

To stop the game, Brad can stop deferring decisions to Angelina. We'll see why that's hard in a moment, because Brad is playing his own game too.

At work, instead of asking for advice, colleagues can say "Don't tell me what to do to help myself, I'll tell you what to do to help me.". Another option is to ask the "helper" to help someone else.

See what you made me do

A game often played alongside "I'm only trying to help you" is "See what you made me do"

When Angelina plays 'I'm Only Trying to Help You', it is then easy for Brad to defer decisions to her. Often this may be done in the guise of considerateness. He may let her decide where to go for dinner or which movie to see. If things turn out well, he can

enjoy them. If not, he can blame her by implying 'You Got Me Into This', a simple variation of "See what you made me do".

Or, like we saw above, he may throw the burden of decisions regarding the children's upbringing on her, while he acts as executor of decisions. If the children get upset, he can play a straight game of "See What you Made me do!".

The payoff here is vindication.

Notice how See what you made me do and I'm only trying to help you are complementary. Both people are playing games, and both games fit together perfectly. Sometimes, these interactions can carry on for years, until they escalate out of control. As Berne puts it, that's when the players end up in jail, the hospital, or the courtroom.

The examples above should've given you a feel for games. I find it interesting to classify games by the kinds of payoff they provide, since games with similar payoffs have similar structure. Understanding this structure makes it possible to quickly notice them in the wild.

Further, knowing the payoff is a very good hint on how to stop the game.

You'll notice that some games have multiple payoffs, in which case, I write their names again, with a line reminding how the payoff shows up.

Again, meet Brad and Angelina. They'll be our key players in these games. Consider this a story of their life, connected via several games. In real life, it's hard to find a couple that plays all these games, but you'll notice how at least some games fit together well.

Vindication

People like to be guilt-free, and would go to insane lengths to rid themselves of guilt. Same for blame, or suspicion. Most people go out of their way to clear their name, because status is involved, and blame and suspicion reduce people's status.

See what you made me do

This game is the poster-child for vindication. We just saw this play out above.

If it weren't for you

Alice is Brad and Angelina's daughter. After an incident at the pool, they don't allow her near a swimming pool. Alice internally fears the water, but she doesn't have to show that anymore. She can, instead, confirm her worldview that she's not afraid, and doesn't go to the swimming pool because her parents don't allow her. "If it weren't for them, I could've swam everyday!". It's not Alice's fault anymore, but her parents. She's vindicated.

Another payoff of this game, as hinted above, is confirming worldviews.

This game can also be played between a couple. An interesting aspect here is why would one side dominate? For example, why would Brad ban Angelina from, say, going out? Berne claims that both sides have a phobia they're avoiding.

"Brad: 'You stay home and take care of the house.'

Angelina: 'If it weren't for you, I could be out having fun.'

At the psychological level the relationship is Child-Child, and quite different.

Brad: 'You must always be here when I get home. I'm terrified of desertion.'

Angelina: 'I will be if you help me avoid phobic situations.'

Debtor

Brad is in debt, and behind payments. He avoids the debt collector. If the debt collector gives up, he wins: Now he has things he didn't have to pay for.

If the debt collector hires a collection agency, Brad knows he has to pay. However, now he can talk about how all debtors are dicks. They could've politely asked for the money back, instead of roughing things up. He can bitch about not paying without losing his social status. He's vindicated.

The game can then morph into a game of "[Why does this always happen to me?](#)"

I'm only trying to help you

Like we saw above: "If my help is not working for you, that's your fault. I'm not guilty. Don't resent me. Try this other thing."

Notice how all the games so far have been 2 player games.

Look how hard I've tried

This is usually a 3 player game.

Angelina is fed up of her husband's antics, and urges him to change to make the marriage work. Secretly, Brad is looking for a divorce. However, he shows up as co-operating. They decide to go to therapy, where Brad plays "Look How Hard I've Tried". He says just enough to show he's co-operating with both, the therapist and the wife. Back home, he's "understanding" for a day or two, after which his behaviour worsens again.

He's effectively forcing Angelina to end the marriage, while saving face himself, by showing every outsider how hard he tried. This is a lot like [playing the role of the husband](#), without actually being one. He's vindicating his position in the marriage.

A more sinister form of the game is “Look how hard I was trying”. Berne writes:

A man is told that he has an ulcer, but keeps it a secret from his wife and friends. He continues working and worrying as hard as ever, and one day he collapses on the job. When his wife is notified, she gets the message instantly: ‘Look How Hard I Was Trying. ‘Now she is supposed to appreciate him as she never has before, and to feel sorry for all the mean things she has said and done in the past. In short, she is now supposed to love him, all previous methods of wooing her having failed.

Self-reprimanding, obtaining forgiveness

Alcoholic

Brad also has a drinking problem. Most people focus on how much he’s been drinking, and how he can’t control himself, and how he should do better. Most nights, he ends up home drunk, hears Angelina complain again and again, and goes to sleep. He wakes up to a splitting headache and a hangover, and Angelina gets him a lemon and honey home-recipe to help.

Berne claims that the drinking is just a side-pleasure. The real payoff comes from the morning-after, or the hangover. That’s where Brad seeks, and obtains forgiveness from Angelina. Being forgiven is the real pleasure that strokes Brad. The cycle repeats as soon as Brad wants another stroke: he’s persecuted at night, when Angelina complains, and then forgiven in the morning, when he has the hangover.

Angelina plays the role of the persecutor and rescuer both.

Berne claims that Alcoholics Anonymous is an organisation that continues the game, but by focusing on converting alcoholics to rescuers for other alcoholics. When a chapter of A.A. runs out of Alcoholics to work on, the members sometimes resume drinking, since there is no other way to continue the game in the absence of people to rescue.

An interesting property of the game is how flexible the roles are: people in one role usually don’t have problems playing the other roles.

Schlemiel

noun INFORMAL • NORTH AMERICAN

a stupid, awkward, or unlucky person.

"he seems like the classic underdog schlemiel"

Schlemiel is very similar to the Alcoholic. Instead of getting drunk himself, in this game Brad creates a mess out of things, like intentionally breaking things, or destroying carpets at a party, and then seeking forgiveness from the host.

Brad's Child is exhilarated because he has enjoyed himself in carrying out these procedures, for all of which he has been forgiven, while [the party host] has made a gratifying display of suffering self-control. Thus both of them profit from an unfortunate situation, and [the party host] is not necessarily anxious to terminate the friendship.

Confirming worldview and stereotypes

We like to confirm what we know, rather than find out what's wrong with our models. These games help reinforce existing worldviews.

Kick Me / Why does this always happen to me?

Brad has a metaphorical board around him that says "Kick Me". When someone kicks, he exclaims incredulously, "Why does this always happen to me?!". It reinforces his worldview that the world is out to get him. For example, when the debt collector hired the collection agency to get him to pay, Brad had been "kicked", thanks to his own action of not paying.

When paired with someone playing "I'm only trying to help you", Brad keeps escalating his behaviour, until the other person is forced to kick them, when Brad can exclaim, "Why does this always happen to me? Everyone's out to get me!"

This might seem similar to Alcoholic, but notice how the payoffs are different. In both cases, Brad instigates, but in Alcoholic, he's seeking to be forgiven, while in Kick Me,

he's seeking to confirm his worldview, so he can maintain status when talking to friends about the idiots in his life.

Debtor

When the debt collector goes hardball, it strengthens Brad's worldview that 'All creditors are grasping'.

A payoff similar to confirming worldviews is reassurance.

Reassurance

Courtroom

This is another 3 player game with a plaintiff, defendant and judge.

Similar to "Look How Hard I've Tried", Brad and Angelina go to therapy again. This time, Brad is annoyed about how his wife broke the sink by using it too much - which forced him to call the plumber and spend a fortune getting it fixed. This annoys Angelina, and she defends herself by telling them that the sink was too old, and, shit happens.

They both turn to the therapist to tell them who's right. Internally, Brad knows he's wrong, but he's seeking reassurance from the outside world to confirm his worldview.

The game element lies in the fact that while the plaintiff is clamouring for victory, fundamentally he believes that he is wrong.

If it weren't for you

"I'm not afraid, father just won't let me swim."

Alice is reassuring herself that she's not afraid.

Why don't you - Yes but

Angelina is at a party, where she's talking about Brad and his new obsession with doing repairs on his own, after a bad experience with a plumber.

Angelina: 'My husband always insists on doing our own repairs, and he never builds anything right.'

Friend 1: 'Why doesn't he take a course in carpentry?'

Angelina: 'Yes, but he doesn't have time.'

Friend 2: 'Why don't you buy him some good tools?'

Angelina: 'Yes, but he doesn't know how to use them.'

Friend 3: 'Why don't you have your building done by a carpenter?'

Angelina: 'Yes, but that would cost too much.'

Friend 4: 'Why don't you just accept what he does the way he does it?'

Angelina: 'Yes, but the whole thing might fall down.'

~~ silence ~~.

Angelina "wins" when all her friends have exhausted their options. Internally, she has demonstrated that they're inadequate at solving her problem.

Berne writes:

"Why don't you - Yes but" is not played for its ostensible purpose (an Adult quest for information or solutions), but to reassure and gratify the Child. A bare transcript may sound Adult, but in the living tissue it can be observed that Angelina presents herself as a Child inadequate to meet the situation; whereupon the others become transformed into sage Parents anxious to dispense their wisdom for her benefit.

Feeling power & Control over others

People like to feel in control, specially over other people.

Now I've got you, you son of a bitch

Brad calls a plumber to get his sinks fixed. Before starting work, they both agree on the full cost. However, the plumber, in the end, added a few dollars extra for some valves. This infuriated Brad, who raged over how the plumber had no ethics and didn't deserve anything in life. Brad didn't pay the bill. In the end, the plumber gave in, and removed the few extra dollars.

Instead of negotiating in a dignified way, Brad took the opportunity to make extensive criticisms of the plumber's whole way of living. On the surface their argument was Adult to Adult, a legitimate business dispute over money. At the psychological level it was Parent to Adult: Brad was exploiting his trivial but socially defensible objection to vent the pent-up furies of many years on his opponent, like his mother might have done in a similar situation.

The plumber here is usually playing a complementary game of "Why does this always happen to me?" - so now he can claim how "All his clients are idiots, whining over a few dollars on a four hundred dollar job."

Berne puts it succinctly, via a game of Poker:

Bob gets an unbeatable hand, such as four aces. At this point, if he is a NIGYSOB player, he is more interested in the fact that Alice is completely at his mercy than he is in good poker or making money.

Kiss Off

Alice and Bob are at a party. Alice signals she's available, makes eye-contact, and Bob strategically tries to place himself closer to her so they can talk. Alice derives pleasure from Bob's pursuit. As soon as Bob commits to talking, and starts a conversation, Alice says "thanks" (or "no thanks") and goes away.

A more sinister form of the game is when Alice/Bob claim they've been assaulted.

The payoff is the feeling of power Alice feels when getting Bob to pursue her and then rejecting him.

Interesting observations

To drive home the idea, here are some interesting insights about games people play.

Complementary games

Some games fit in well with other games, where both people can be playing their own game, and both people win.

- Kick me <--> Now I've got you, you son of a bitch <-later-> Ain't it Awful
- Kick me <--> rapo
- Alcoholic <--> I'm only trying to help you
- Why don't you - yes but <--> I'm only trying to help you

Reversible roles

Players playing a game are adept at all roles in the game, and can switch. For example: When playing "Why Don't You .. Yes But":

Sometimes you ask the questions and make suggestions (and perhaps play "I'm only trying to help you"). Other times you're posing the problem and expecting to hear suggestions you've already thought of yourself.

Rough model

Almost every game seems to consist of a victim and a persecutor. Remember, transactional analysis dictates we can't have two persecutors! That's just two angry people talking over each other. For communication to continue, one needs to be the "victim".

Sometimes, it's the persecutors game, where they're enjoying the control, or seeking justification for their actions.

Some other times, it's the victims game, trying to confirm their worldview, self-reprimanding, or seeking forgiveness.

Other times, it's a complementary game played by both, the persecutor and the victim.

In 3 handed games, the third role is usually the rescuer.

Children are impressionable

.. and pick up games from their parents. This is what builds up their Parent Ego state. If as they grow, their Adult state doesn't get strong enough, in times of crisis they'll respond exactly as their parents did, which may not be what you want.

Sabotage

This model makes sabotage legible.

I'd get frustrated when I saw people on the cusp of breaking out of a bad situation, but messing things up. I realised the game wasn't to break out, but to maintain status quo. To continue the game, people will sabotage themselves.

In its more sinister form, it's when others sabotage you to continue their game. Berne writes:

“Even the close relatives of the patient who complained most loudly about the inconveniences caused by his infirmity, may eventually turn on the therapist if the patient makes definitive progress. [...] All the people who were playing ‘I’m Only Trying to Help You’ are threatened by the impending disruption of the game if the patient shows signs of striking out on his own, and sometimes they use almost incredible measures to terminate the treatment.”

Antithesis: How to Stop Games

With the games in hand, and a feel for their destructive power, the next natural question is how do you stop them?

Denying the payoff stops the game.

Sometimes, it may be hard to figure out what game is being played. Since the ultimate payoff is what's driving it anyway, testing a few different kinds of responses can help you figure out the game.

If they react more intensely, or stop doing what they're doing, you're on the right track. It might also help to understand which position the game is being played from. For example, if it's Parent-to-Child, responding like an Adult breaks communication. This is a bit hard to do, since we're conditioned to respond in a way that continues communication, which is exactly what the player wants, since it allows the game to continue.

I'll go payoff-by-payoff, and choose one, or a few, representative games, and discuss ways to stop them.

See What You Made Me Do - Vindication

Since Brad is deferring responsibility and vindicating himself, the best way for Angelina to stop this game is to flip the questions back to Brad. "It's your choice, sweetheart. We'll do whatever you want to do". Of course, if this turns into a game of *See What You Made Me Do* with Angelina as the lead, then they're in a loop that would need a third agent to fix.

Look How Hard I'm Trying - Vindication

In the marital case, stopping the game means noticing that Brad isn't genuinely involved in the therapy sessions, and dismissing him from therapy. This destroys Brad's claim that he's been trying hard. He can still get a divorce, but not with the social psychological advantage of "look how hard I tried".

With kids playing look how hard I'm trying, it's a good idea to explain how responsibility works, the game, and how they may be judged on outcomes, not what they do. You're explaining how trying hard doesn't vindicate them.

Self-reprimanding & Obtaining Forgiveness

In both Alcoholic and Schlemiel, since the players are looking for forgiveness, the best way to stop the game is to tell them they won't be forgiven, and don't forgive them.

Berne writes:

“Anti-‘Schlemiel’ is played by not offering the demanded absolution. After [Brad] says ‘I’m sorry’, [the party host] instead of muttering ‘It’s okay’, says, ‘Tonight you can embarrass my wife, ruin the furniture and wreck the rug, but please don’t say “I’m sorry”.’ Here [the party host] switches from being a forgiving Parent to being an objective Adult who takes the full responsibility for having invited [Brad] in the first place.”

Why don't you ... Yes but - Reassurance

The best way to stop reassurance games is to not play your part. When someone talks about how they're suffering, do not play "I'm only trying to help you". Don't give suggestions. Instead, flip the question back to them: "So, what are you going to do about it?"

It's a similar story with "If it weren't for you". Instead of banning them from doing X, let them know they're free to do X. But that means getting rid of *your* insecurities related to X.

Sometimes, when this happens, the players may escalate the game. Berne recommends going to therapy, as resolving deep ingrained games should best be left to professionals.

Control and Power over others

The best way to stop games like "Now I've got you, you son of a bitch", is to notice them, make the contract explicit, and stick to the contract no matter what. These players are looking for opportunities for you to slip up.

Of course, if possible, don't do business with them.

You've got to be careful about not just the contract, but everything around it:

In everyday life, business dealings with NIGYSOB players are always calculated risks. The wife of such a person should be treated with polite correctness, and even the mildest flirtations, gallantries or slights should be avoided, especially if the husband himself seems to encourage them.

Epilogue

When you go about life without questioning something, things "just happen". There's nothing weird about it, it's "just how things are". When someone starts questioning this random something, and comes up with a model to explain things, it's a new lens you can't unsee. Suddenly, things simply fit in!

The Theory of Games People Play is one such model. People don't "just do things", there are incentives and psychological motivations that haven't been made clear to you.

Of course, this model isn't perfect, and doesn't explain everything. It's [one of many perspectives](#). But, it's how the journey to building [better models](#) begins. If you never get introduced to models, you never begin. Once you have a model for social interactions, you can notice where it breaks down, and look for other models to fill in the gap.

1. Affiliate link
2. Note the capitalisation. When I say parent, I mean the person who is the parent. When I say Parent, I mean the Parent ego state. Same for Child and Adult.
3. It's interesting to see how pick up artists figured out these same games, but have different names for them. For example, in this case, they "leave her alone".
4. Not all games are bad! There are some good games, too. Annoyingly, they usually don't show up in therapy sessions, so study on them is limited. ↩
5. Usually. Other times, for people deeply entrenched in games, it escalates their actions till they get the payoff.

Gödel, Escher, Bach by Douglas R. Hofstadter

Gödel, Escher, Bach, like this sentence, has two parts: the first cleverly makes its point while engaging the reader with playful use of self-reference; the second is less interesting, slower developing, and far longer ... in fact, it's so long that you're likely to quit reading it before it ever gets to its

Part I is an exposition of many interesting and deeply related ideas: formal systems like math and physics acquire meaning by modeling the world; recursion gives these systems power but also opens the door to self-reference; and self-reference ultimately poses a serious problem for these systems. These ideas build to the statement and proof of Gödel's Incompleteness Theorem.

Part II, roughly speaking, claims that the ideas of part I have something to do with artificial intelligence and the nature of consciousness. This is a review of GEB part I, though I'll briefly touch on part II at the end.

Before I start, let me tell you some things that *won't* be in this review because you really can't get them from anywhere but GEB itself.

First, GEB author Douglas Hofstadter sees interconnections *everywhere*. GEB initially comes across as a perplexingly well-regarded conspiracy theory text. But reading on, you come to see the magic: all of the conspiracies are *actually true*. This is reflected in the title of the book: Hofstadter first set out to write a pamphlet on Gödel's theorem, but as the pamphlet grew into a book and then a 700-page tome he came to believe that the work of Escher and Bach were too interconnected with the subject matter to be omitted. So if you want to know why Gödel numbering is just like DNA transcription, or recursive transition networks are just like renormalization of elementary particles, sorry. You'll have to check out GEB from your local library.

Second, this review will feature very few of Hofstadter's actual words. The reason is simple: there's way too many of them. In a previous draft of this review, I tried quoting out of GEB for a few simple things, but it would always turn out like "Hofstadter thinks humans are sometimes different than machines: [300 word quote that somehow essentially involves an analogy about how *you* think your wife wants you to turn off the TV, but *she* wants you to start a government-overthrowing revolution] (page 37)." Though I personally enjoy the colloquial, meandering feel of Hofstadter's writing, many find it infuriating, and it's probably a key reason that the few people who attempt reviews of GEB tend to quote Lewis Carroll as much as Hofstadter himself.

And third, GEB is *really* idiosyncratic in a way no one can imitate. The book's chapters are each separated by an entertaining Carrollian dialogue featuring the quick-thinking tortoise and his slower friend Achilles; these dialogues illustrate key ideas that reappear later in the text, imitating the way themes reappear in Bach's fugues. Hofstadter has an axe to grind with Zen

Buddhism, and the first application of a formal logical system he develops in the text is to refute a Zen koan about grinding axes. He also enjoys taking pot shots at composer John Cage for basically no reason.

Overall, I think GEB is a really good book. In fact, I insist that it's better than you expect [even after taking my insistence into account](#). Rationalist caliph and LessWrong founder Eliezer Yudkowsky, on whom GEB was an early influence, agrees:

"*Gödel, Escher, Bach* by Douglas R. Hofstadter is the most awesome book that I have ever read. If there is one book that emphasizes the tragedy of Death, it is this book, because it's terrible that so many people have died without reading it."

So, lest you die without familiarizing yourself with GEB, let's get started.

I. Formal systems and interpretations

The basic object of study in GEB is what Hofstadter calls a *formal system*. A formal system consists of:

- A collection of allowable characters out of which we can form strings (sequences of characters)
- A collection of strings called "axioms"
- A collection of rules, or "inference rules," for changing some strings into others

Huh? Let's start with a simple, meaningless example called the MIU-system.

The MIU-system:

- Allowable characters: M, I, and U. (So strings are things like M, UMM, MIMMIUM, UMIIMUMUIMIM, etc.)
- Axioms: MI
- Rules:
 - Rule I: given a string that ends in an I, you can add a U to the end.
Example: from UMI, form UMIU
 - Rule II: given a string of the form Mx where x consists of M's, I's, and U's, you can form the string Mxx
Example: from MIU, form MIUIU
 - Rule III: given any string with III appearing somewhere inside, you may replace III with U
Example: from MIII, you can form MUI (by replacing the middle III with U). You can also form MIU (by replacing the ending III with U).
 - Rule IV: given any string with UU appearing inside, you may delete UU
Example: from MUUI, form MI

Let's call a string a *theorem* if you can produce it from the axiom MI using the inference rules. For example, I claim that MUIIU is a theorem; in support of this I offer the following "proof":

- (1) MI (axiom)
- (2) MII (using rule II)
- (3) MIII (using rule II)
- (4) MIIIIU (using rule I)
- (5) MUIU (using rule III)
- (6) MUIUUIU (using rule II)
- (7) MUIIU (using rule IV)

There you have it – MUIIU is a theorem (as are all of the strings obtained along the way).

Hold up, axioms? theorems? Readers who've seen some mathematical logic might know where I'm going with this.

The terminology is chosen to suggest the following. We imagine that the given rules are "rules of logical inference," analogous to rules in classical logic like "if you know 'P' and you know 'if P then Q,' then you may conclude 'Q.'" We imagine that the strings of our system are logical statements written in some formal language. And we imagine that the axioms are some logical statements that we declare to be true. So the "proof" above is akin to starting from a known axiom and using the rules of logical inference to deduce some desired theorem, sorta like a proof! Formal systems are a way of mechanistically codifying logic; even a beginning programmer could write a program that starts from the axioms and generates a list of theorems by applying the rules. In fact, this is the basic principle behind how automated theorem-provers like Coq work.

After introducing the MIU-system, Hofstadter offers the following puzzle, which I pass on to you: is MU a theorem? Try to figure it out yourself if you'd like, or read on to find the answer later.

In this example, the MIU-system doesn't seem to reflect the structure of anything we would care about. In contrast, the next example-and-half do: they are meant to model multiplication of natural numbers.

The tq-system:

- Allowable characters: t, q, -
- Axiom: -t-q-
- Rules:
 - Rule I: given a string xtyqz where x,y,z are strings consisting of only hyphens, you can form x-tyqzy
 - Rule II: given a string xtyqz where x,y,z are strings consisting of only hyphens, you can form xty-qzx

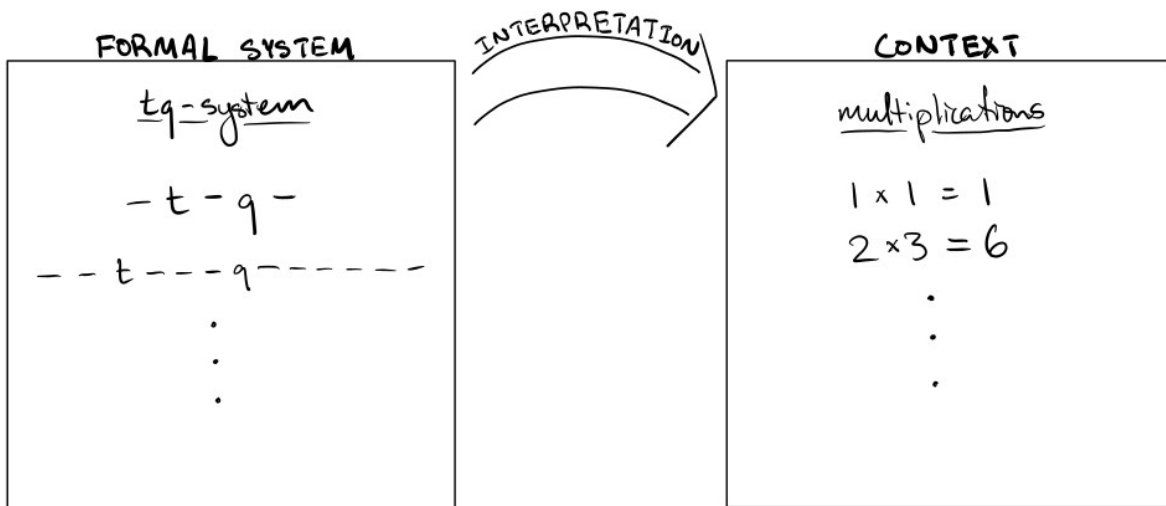
Unlike the MIU-system, the tq-system comes with an *interpretation* which converts strings of the formal system into meaningful statements in some *context*. In this case, the context is “multiplications,” and the interpretation looks like

- t ⇒ times
- q ⇒ equals
- ⇒ one
- ⇒ two

and so on. This interpretation turns the axiom -t-q- of the tq-system into the multiplication “one times one equals one” and the theorem --t---q----- (proved below) into the multiplication “two times three equals six.”

Proof:

- (1) -t-q- (axiom)
- (2) --t-q-- (rule I)
- (3) --t-q---- (rule II)
- (4) --t---q----- (rule II)



We can think of an interpretation as giving *meaning* to a formal system. Uninterpreted, --t---q----- is a meaningless string of characters, same as the strings of the MU-system. But equipped with the interpretation above, this string comes to *mean* the multiplication “two times three equals six.” An analogy: to a child ignorant of the world, a globe is just a meaningless spinning toy. But once the child learns that pictures on the globe (the formal system) represent (interpret to) masses of land on the actual Earth (the context), aspects of the globe start to carry meaning – the fact that the splotch of green labeled “Asia” is larger than the one labeled “Australia” corresponds to the continent Asia having a larger land-mass than the continent Australia.



Liberation, by M.C. Escher. Strings in formal systems (the triangles at the bottom) are transformed into meaningful statements (the birds) via interpretation.

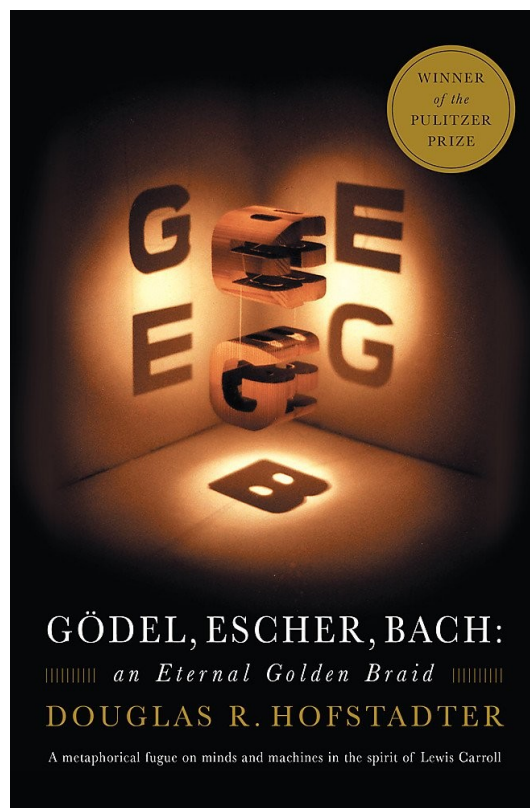
At this point, three caveats are in order.

First, you should not think that a formal system necessarily has only one interpretation. For example, here's another interpretation of the tq-system, now into the context of divisions:

t \Rightarrow equals
q \Rightarrow divided into
- \Rightarrow one
-- \Rightarrow two

and so on, so that --t---q----- now interprets to "two equals three divided into six." In a case like this, it'd be a mistake to argue about what the "true meaning" of the string --t---q----- is; the

correct takeaway is that both meanings are encoded simultaneously. Even this simple example of a double-entendre is somewhat interesting: it demonstrates that the structure of multiplications is “the same” as the structure of divisions (borrowing a word from mathematics, Hofstadter would say that multiplications and divisions are “isomorphic”).

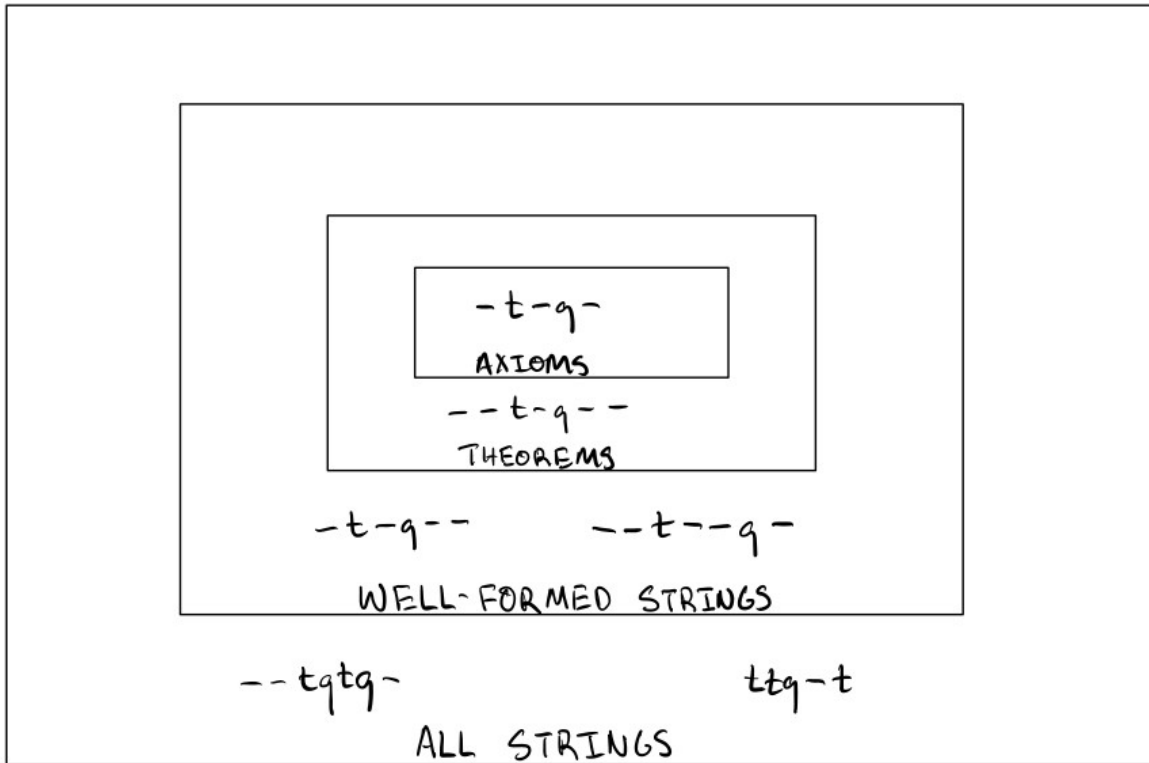


The cover art is a real photograph of two carved blocks of wood. Depending on which interpretation (angle of the light) you use, you can pull three different meanings out of each block.

Second, not all strings of the tq-system come out meaningful under interpretation. The tq-system also contains strings like ttq-t which don't correspond to any multiplication. Let's call a string *well-formed* if it does carry meaning under our choice of interpretation. This includes strings like -t-q-- which do mean something (one times one equals two) even though that something is false.

And third, all of the *theorems* of the tq-system are not only well-formed, but they also represent true multiplications. For example the theorems -t-q- and --t---q----- interpret to the true multiplications "one times one equals one" and "two times three equals six." (The well-formed string -t-q-- doesn't, but that's fine because it's not a theorem.) This is really important, so let's make it a requirement: if I call something an "interpretation" of a formal system, **I will always mean that the theorems are well-formed and come out true** under the interpretation.

For a counterexample, if we changed '-' to mean "two," then we wouldn't have an interpretation anymore since the theorem -t-q- would represent the multiplication "two times two equals two," which isn't two – achem excuse me – true.



As a final half-example of a formal system, let's augment the tq-system so it can prove theorems representing statements like "6 is composite."

The tqCP-system:

- Allowable characters: t,q,-,C,P
- Axioms: same as tq-system
- Rules: same as the tq-system, plus
 - Rule III: given a string $xtyqz$ where x,y,z consist of at least two hyphens, you can form Cz

The interpretation I intend for the tqCP-system into the context of "arithmetical statements" looks the same as the tq-system, plus:

$Cx \Rightarrow x$ is composite

$Px \Rightarrow x$ is not composite (or equivalently, x is prime)

What's up with having a P when the inference rules don't allow it to appear in theorems? More on that later.

II. Jumping out of the system

I claimed above that the given interpretation of the tq-system was valid, i.e. that it transforms theorems of the system into true multiplications. How do I know that? Sure I gave two examples, the theorems $-t-q-$ and $--t---q-----$, but how can I be sure that every one of the infinitely many theorems of the tq-system interpret to true multiplications?

I'll argue like this. First of all, the axiom $-t-q-$ interprets to a true multiplication (one times one equals one). Second, we note that given a string $xtyqz$ which represents a true multiplication (x times y equals z), rule I produces a string which represents a true multiplication ($(x + 1)$ times y equals $z + y$). Same goes for rule II. As our axioms are true and our rules of inference preserve truth, all of our theorems must be true as well!

Where did the reasoning in the last paragraph take place? It certainly wasn't a proof "inside the tq-system," since those proofs just look like lists of tq-strings which obey the inference rules. Rather, it was an example of "stepping outside of the system." We reasoned *about* the tq-system using ordinary reasoning, not the internal formal logic of the tq-system. After all, the system knows nothing about the interpretation we've given it – it doesn't know that its theorems are supposed to represent multiplications. So we can't possibly hope to prove the validity of the interpretation by working within the tq-system. We had to step outside.

Here's another example of stepping outside the system. We just saw that every theorem of the tq-system represents a true multiplication. In fact, the converse is also true, namely that every true multiplication is represented by a theorem of the tq-system! If you're interested, you may wish to prove this – it will require stepping outside the system. Then, using this observation, you can derive theorems of the tq-system "from the outside." For example, since $---t---q-----$ represents a true multiplication, we know that it must be a tq-theorem. Again, this isn't a "proof" in the formal sense, because a proof is a sequence of tq-strings produced by applying rules. It is a proof from the outside.

Hofstadter points out that jumping outside the system is an important feature of intelligence. Before I introduced the tq-system I told you what my intended interpretation was. But even had I not, it's very likely you would have discovered it after a few minutes writing down tq-theorems. Instead of mindlessly churning out an ever longer list of theorems, you would instead gradually notice the patterns, put down your pencil to think, and discover that you can predict what all the tq-theorems are without writing them down. These are all outside-the-system activities.

Even now, you're likely making frequent jumps out of your "reading this book review" system. Perhaps you're pausing to check if you're thirsty or need to go to the bathroom. And perhaps now you're asking yourself if it counts as jumping out of the system if I just told you to do it. And

maybe you're now trying to do something I didn't tell you to do just to prove that you really can jump out of the system. (sorry)

Contrast this with the behavior of a graphing calculator running a basic program that prints out a list of tq-theorems. The graphing calculator will never stop executing its code, step back to survey the data, notice the pattern, and print out IT'S THE MULTIPLICATIONS YOU DUMMY. Of course a human is ultimately some program, albeit a very complicated one running on the most powerful computer in the known universe. Accordingly, there is some system out of which we are unable to step, the same way biological evolution is unable to step back, take a look at the data, and shout into the void [JUST KEEP MAKING MORE CRABS](#) YOU DUMMY. The point isn't that human intelligence is "special" in some way that purely mechanistic reasoning can never replicate. The point is simpler: intelligent systems seem to be able to identify and run subtasks, as well as to monitor these subtasks from the outside and determine when to stop doing them.

III. Truth vs. provability

“ ‘Snow is white’ is true if and only if snow is white.”
– Alfred Tarski

At this point, it's possible you're mixing up the notions of truth and provability. If so, don't feel bad: so did literally everyone for the whole history of logic until 1930. That's when German logician Kurt Gödel announced his namesake *Incompleteness Theorem*, a consequence of which is that truth and provability really must be considered as separate notions. A goal of GEB part I, and of this book review, is to outline the key ideas going into this theorem. In this section I'll explain the distinction between truth and provability and state Gödel's theorem. But first, a story about Gödel.

Life in late 1930s Europe wasn't treating Gödel well. For one, he was unable to find a position because he had too many Jewish friends (a common side-effect of being a mathematician). And to make matters worse he had been conscripted to the German army. So Gödel did the logical thing: he fled to the U.S., got a position at Princeton, and hung out with his buddy Albert Einstein (who confessed that the only reason he showed up to work was for "the privilege of walking home with Gödel.") While studying for his U.S. citizenship exam, Gödel claimed to have discovered this one weird trick for legally turning the U.S. into a dictatorship (anti-fascists hate him!) Despite Einstein's warnings to Definitely Not Bring That Up, Gödel totally Brought It Up during his citizenship interview. Fortunately, Einstein, who was there serving as a witness and also knew the interviewer, managed to smooth everything over. Gödel became a citizen. I'm not sure what the moral is, but hopefully this gives you a taste of the mind of Kurt Gödel.

Okay, back to the logic textbook masquerading as a book review. A good way of thinking about the truth/provability distinction is that provability comes from the formal system and truth comes from the interpretation+context.

Provability is simpler, so let's tackle it first. Calling a string in a formal system *provable* is just a fancy way of calling it a theorem. That is, "provable string" and "theorem" are synonyms. This should make sense: remember that "theorem" just means something you can deduce from the axioms using the inference rules, i.e. something you can "prove." For example, the strings $-t-q-$ and $C-----$ are provable in the tqCP-system, but $-t-q--$ is not. In the MIU-system, MI and MUIIU are provable but (spoiler!) MU is not. Note that provability is a purely formal notion, i.e. it depends only on the formal system and not on whatever interpretation you attach to it.

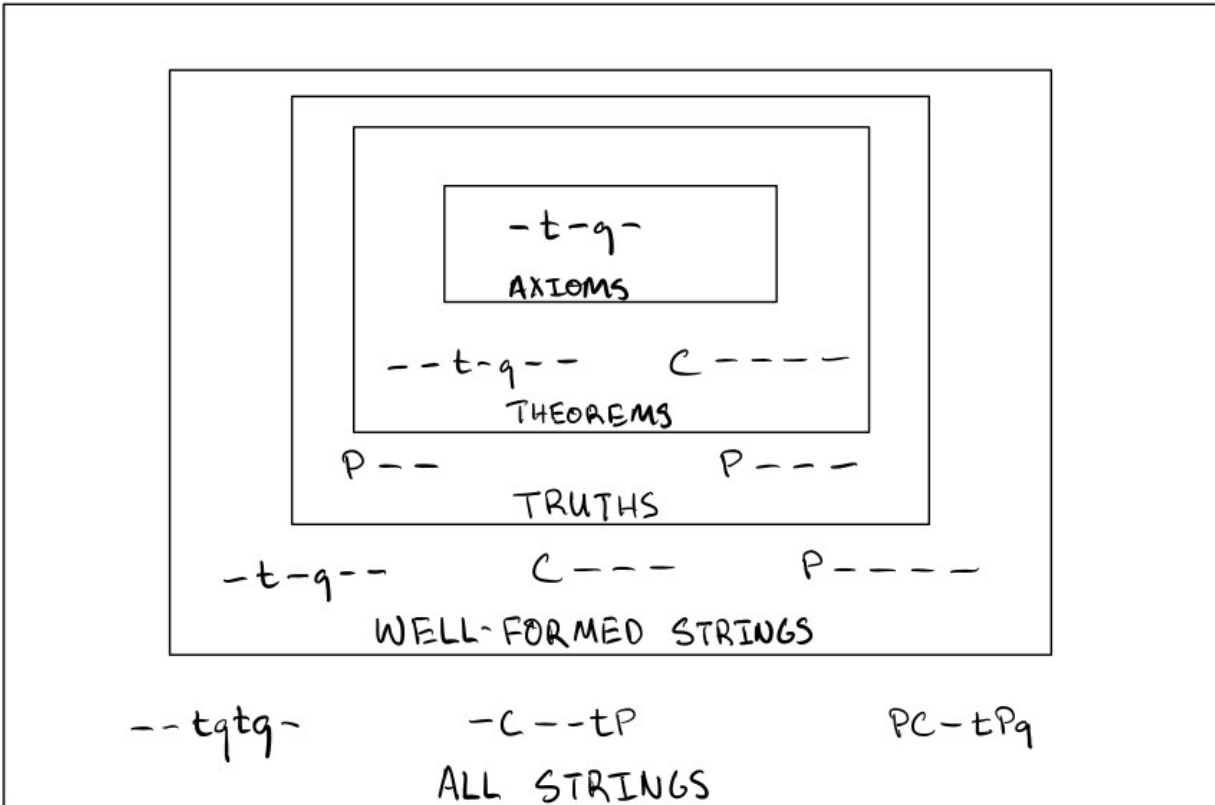
Truth on the other hand relies on a choice of interpretation. Given a formal system with an interpretation, we say that a string of the system is *true* if it comes out true under the given interpretation. For example, $--t---q-----$ is true because two times three does equal six, but $P----$ is false because four isn't prime. We can't say whether MIII or MMU are true because we don't have an interpretation for the MU-system in mind.

Since by fiat all of our interpretations translate theorems to true statements, we know:

in a formal system with an interpretation, all provable strings of the system are also true.

Or more succinctly: if provable then true. This is really important: it's why mathematicians and physicists can write some funny little squiggles on paper, do logic on them, produce some different funny squiggles, and be confident that the new squiggles actually mean something true about the universe!

You might be tempted to believe the converse: that every true statement in a formal system is also provable. (Or at least, you might have been tempted to think that if I didn't have a whole section titled "truth vs. provability".) But consider the string $P--$ of the tqCP-system, which interprets to "two is prime." This string is certainly true, since two is prime. But it is *not* provable in the tqCP-system – in fact, none of the rules of the system allow you to produce a theorem with the character P.



You're probably thinking that this demonstrates that the tqCP-system is *bad* in some way, or at least woefully *incomplete*. Perhaps you're tempted to augment the tqCP-system by adding a new rule: if Cx is not a theorem for some x consisting of only hyphens, then Px is a theorem. But there's an issue here: applying this rule requires making a list of all (infinitely many) theorems of the tqCP-system and checking that Cx is not among them. But this is not the sort of simple, mechanistic rule that our formal systems are allowed to have – no person, and certainly no computer, could ever finish writing down all the theorems and checking that $C--$ is not among them. You might be able to prove *from outside the system* that $C--$ is not a theorem, but such "outside the system" reasoning has no bearing on provability *inside* the system.

Fear not: Hofstadter does explain a way of augmenting the tqCP-system to be able to prove statements like $P--$ (though it requires adding new characters as well as new rules). So *now* can we admit that the tqCP-system is no-good, and that we should root out all formal systems that can't prove all their truths and replace them with ones that can?

If only! Gödel doesn't tolerate citizenship interviewers ignorant about weird tricks for making dictatorships, and he won't tolerate our nonsense either. Here is his theorem.

Gödel's Incompleteness Theorem: any sufficiently rich formal system (with an interpretation) has strings which are true but unprovable.

A formal system which *is* able to prove all of its truths is called "complete." So Gödel's theorem says that every sufficiently rich formal system is incomplete – there will always be unprovable truths.

V. Self-reference and the proof of Gödel's theorem

Suppose that Donald Trump were to walk into your room and say:

"This sentence is a lie."

It being Donald Trump, you might suspect that he's lying. But if that were the case, then "This sentence is a lie" would be the truth, so Donald would be telling the truth ... a contradiction! Likewise, if you give him the benefit of the doubt and suppose he's telling the truth, you'll come to find that he's lying, another contradiction (and another Tuesday in politics).

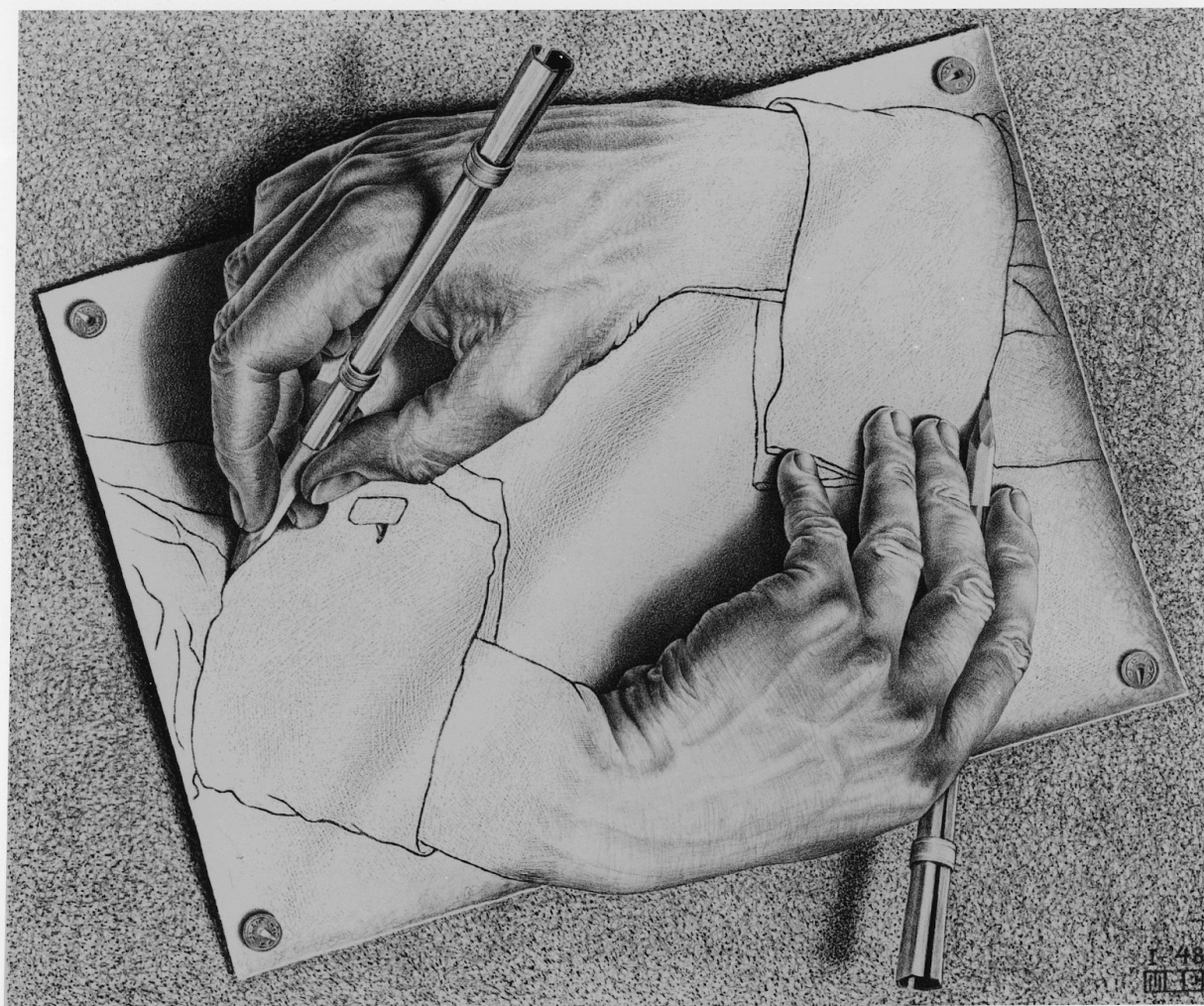
This is called the liar's paradox, and it's the basic idea behind the proof of Gödel's theorem. The core of the issue is that we have a system (the English language) trying to model itself, and we've exhibited a sentence whose interpreted meaning references that very same sentence. This snake-eating-its-own-tail pathology can be arranged to create other [similar paradoxes](#).

You might think that we can fix things like this with a simple rule like "no interpretation of a formal system can have the context be that very same system." Unfortunately, things aren't so easy. Consider the following two-step version of the liar's paradox.

The German sentence below is false.

Der obige englische Satz ist wahr. ("The English sentence above is true.")

Here, "sentences in English" has an interpretation with context "sentences in German." But "sentences in German" itself can model "sentences in English." And although each sentence by itself is perfectly harmless, the whole is paradoxical!



Drawing Hands, by M.C. Escher, an illustration of the two-step liar's paradox.

Part of the issue is that English is *too rich*. That is, it's able to talk about concepts like "truth" and "falseness" as well as support self-reference. It's also rich enough to model systems (like German) which are themselves rich enough to model English, enabling the two-step liar's paradox. These aren't issues that are easy to patch; it's not clear how much of English we would need to remove to make it "not too expressive." Perhaps in doing so, we would destroy our ability to say anything useful.

English is too fuzzy to work with, so instead Gödel works with statements of number theory – things like "two plus two equals four" and "four is a factor of eight." It ends up that while number theory isn't expressive enough to talk about the *truth* of number-theoretical statements, it is expressive enough to talk about the *provability* of number-theoretical statements.

The idea of Gödel's proof is to encode a "provability version" of the liar's paradox into number theory. That is, given a formal system rich enough to model number theory, Gödel comes up with a string G of the system whose meaning under interpretation is:

“G is not provable.”

If G were false, then G would be provable and hence true, a contradiction. So G must be true, making it an unprovable truth. It follows that the formal system in question is incomplete.

The rest of this section fleshes out this idea in more detail, using an idea called Gödel numbering. I think it's pretty cool, but if it's not your cup of tea, feel free to skip to part VI.

As a warm up, recall the MU puzzle from above: determine whether MU is a theorem of the MIU-system. I will now demonstrate that the answer is "no – MU is not a theorem." The idea is to encode "MU is a theorem of the MIU-system" as a claim about number theory, and then figure out if that claim about number theory is true.

To do this, let's first transform strings of the MIU-system into numbers by the rule:

M \Rightarrow 3

I \Rightarrow 1

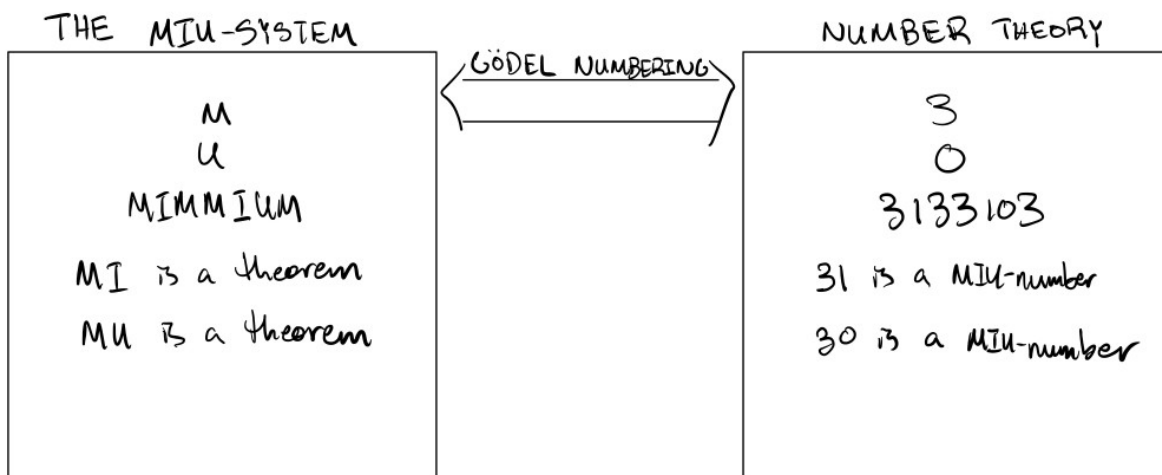
U \Rightarrow 0

For example, MIUII is the number 31001, MU is the number 30, and the axiom MI is the number 31. Under this transformation, the rules of the MU-system can be stated arithmetically. For example, rule I says that if a number has units digit 1, then you may multiply it by 10 (thereby appending a 0 to the end). Or more formally:

given a number of the form $10m + 1$, you may form the number $10*(10m + 1)$.

You can do the same thing for the other rules too.

Let's call a number which corresponds to a theorem of the MIU-system a MIU-number. So we've transformed the claim "MU is a theorem of the MIU-system" to the equivalent claim "30 is a MIU-number," which can also be stated as "30 can be formed from 31 by repeatedly applying such-and-such arithmetical operations." This might not seem like progress, but it is! The claim "30 is a MIU-number" is a *number theoretical statement* (though perhaps not an interesting one). In essence, it's similar to – but more complicated than – the more familiar statement "216 is a power of 6" i.e. "216 can be formed from 1 by repeatedly applying the multiply-by-6 operation."



Now we can dispose of the MU puzzle by proving a proposition about MIU-numbers:

No MIU-number is divisible by 3.

I'll leave the proof to you – it's not hard, especially if you remember [the rule](#) for checking whether a number is divisible by 3.

Since MU corresponds to 30, which *is* divisible by 3, we deduce that 30 is not a MIU-number. Hence MU is not a theorem of the MIU-system, and we're done. If you're rightly baffled, you can press pause on your computer screen now to reflect deeply on what's happened. You can press play when you're ready to resume the proof of Gödel's theorem.

The procedure above turned strings of the MIU-system into numbers, and claims about those strings into statements of number theory. This is called *Gödel numbering*, and it can be done to any formal system. Via Gödel numbering, the claim "MU is a theorem" *about* the MIU-system corresponds to the number-theoretical claim "30 is a MIU-number." In other words, despite the MIU-system having no interpretation that gives its strings meaning, Gödel numbering gives number-theoretical meaning to certain claims *about* the MIU-system.

Could something interesting happen if we Gödel number a system that *already has* an interpretation into number theory? Could the meaning acquired through the interpretation clash with the meaning induced by Gödel numbering? Could my rhetorical questions get any more leading? Is the answer to all of these yes?

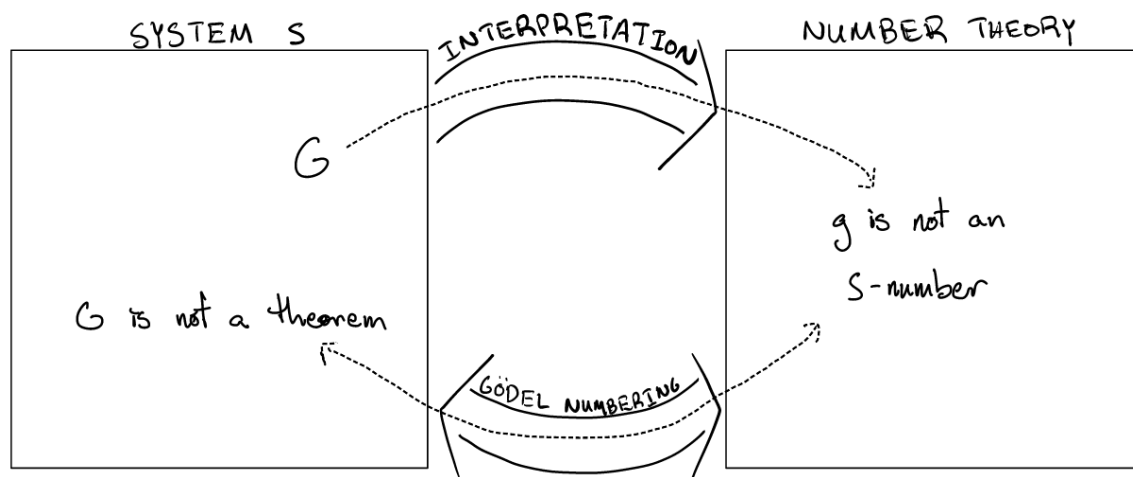
In GEB, Hofstadter spends two chapters constructing an explicit example of a formal system that models number theory, called Typographical Number Theory, or TNT (foreshadowing that it will blow itself up). For the sake of concreteness, he then proves the Incompleteness Theorem for the system TNT. Nevertheless, the same proof works for a general formal system *S* with an interpretation into number theory, and I'll explain it here in this more general language.

(here comes the technical meat; please set your brains to "think very hard")

Suppose we are given a formal system S with an interpretation into number theory. And suppose that the formal system is "rich enough" in the sense that any statement about number theory can be rendered as a string of S . We want to show that S has an unprovable truth. Fix a Gödel numbering for S , i.e. a correspondence between characters of S and digits which turns all the strings of S into numbers and all the rules of S into arithmetical rules. As before, let's call a number an S -number if it corresponds to a theorem of S .

Given a string G of the system S , let g be the number corresponding to G under the Gödel numbering. Now, " G is not a theorem of S " is equivalent to the number-theoretical claim " g is not an S -number." But the number theoretical claim " g is not an S -number" can in turn be rendered as a string of S (as can any number theoretical claim, by assumption). Let's call this string G' .

In a situation like this, Gödel gave a magic recipe (or see chapters 13 and 14 of GEB) for cooking up a specific string G such that the resulting G' is the same as G . Thus, this G interprets to the statement " g is not an S -number," which is true if and only if G is not a theorem of S . Informally, we might say that G carries the meaning " G is not provable in S ." And now we're done: if G is false, then G is a theorem of S , and is therefore true, a contradiction. So G is true, and thus G is not provable. Thus G is an unprovable truth and S is incomplete. Q.E.D.



I'll end this section with an exercise for those interested: how is this proof like the proof of undecidability of the halting problem? (For solutions, please consult *Gödel, Escher, Bach* by Douglas Hofstadter.)

VI. GEB Part II

As a math grad student, I'm not a bad person to write a book review of GEB part I. On the other hand, I'm vastly unqualified to say anything about GEB part II. Buuuut I'll say a bit anyway.

My executive summary of GEB part II is: you know all those crazy ideas about self-reference, meaning, etc. in part I? Those all have something to do with intelligence and consciousness.

Obviously this summary isn't doing it justice. There's a bunch of big ideas in it, a slew of interesting paradigms, and my overall sense is that most of them are wrong.

Why so? Part of it is that Hofstadter argues from first principles for a bunch of claims that don't seem to have aged well since GEB's 1979 publication. For example:

1. While he avoids positing the existence of a so-called "grandmother neuron" – that is, a neuron whose sole job is to fire whenever you need to make use of the concept "grandmother" – he does seem to think that something kinda like this is true: that there is a "grandmother module" in the brain – perhaps a collection of neurons – which activates whenever you think of a grandmother.
2. He seems to believe that the way we think thoughts is for all of our various modules to fire together in roughly the same way that a bunch of words are said together to form a sentence. E.g. the thought "My grandmother is happy" boils down to the modules in your brain representing "grandmother" and "happy" firing together, along with some additional information to specify that it is "my grandmother" instead of just "a grandmother" and things like that.
3. His paradigm of (artificial) intelligence seems to involve intelligent systems working symbolically within a formal system while simultaneously connecting the patterns of the formal system to whatever problems it's trying to solve.

Idea 1 ... well it's actually not so bad. We have a decent understanding of how the brain does the most basic steps of visual processing – things like edge detection – and we've identified some particular neurons that fire together in certain ways to encode information about the orientations of lines. This is kinda like a low-level version of grandmother module for very simple concepts like vertical-ness (though apparently [predictive processing](#) has another take on what information exactly is being represented). Also, some people with brain damage have [trouble distinguishing](#) among inanimate objects but not among animate ones, which if you squint kinda looks like evidence we have an "inanimate object" module in our brain which sometimes gets damaged.

On the other hand, my girlfriend studies the brain of *C. elegans*, a nematode with 302 neurons (the human brain has around 86 billion neurons). She specifically studies olfaction – how the worm processes smells. You would think if there were a small module representing the concept of sulfuric-odor-ness or whatever, we would have found it by now. Instead, it doesn't really look like the information of the smell is stored in any one place. Her advisor – a fancy professor who

knows a whole lot about the *C. elegans* brain – thinks that the whole brain is involved even in simple things like detecting a smell. And my sense of the more complicated field of human neuroscience is that it's not so keen on the “such-and-such part of the brain represents such-and-such concept” paradigm either.

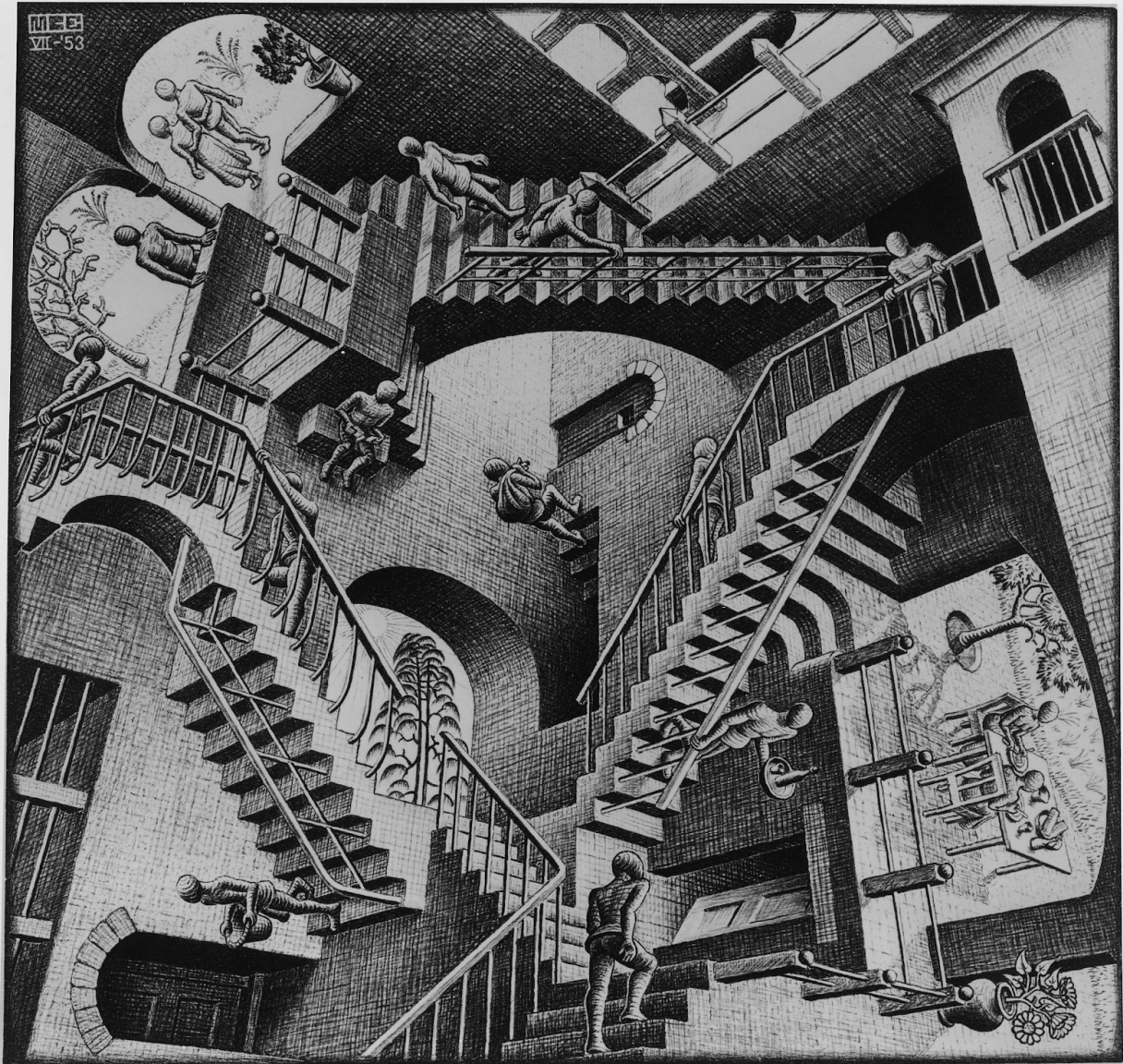
Given these issues with idea 1, idea 2 looks only more ludicrous. I'm not even sure what to say here other than to gesture generally at everything I know about the brain (warning: not much) and note that none of it looks at all like this paradigm.

Maybe idea 3 – building an AI that constructs formal systems to model the world, and then works within these formal systems to generate new data/predictions/whatever – could work. But it seems so, so far from the direction that AI has actually gone. I don't see how this paradigm could survive a collision with GPT-3, which can write ([and draw](#)) better than most people I know, but fails at multiplying large numbers. If GPT-3 secretly works by constructing a super complex formal system that models the human-produced text in its training data, how come it couldn't come up with a much simpler formal system (like the tq-system) for modeling the multiplications?

Actually, why am I even talking about GPT-3? Are we really supposed to believe that the human brain is able to model all the complexities of language, society, whatever by computing really rapidly in hyper-complex formal systems ... yet I can't consistently subtract 6 from 11? Whatever my brain is doing to model the world, constructing and working in formal systems doesn't seem like one of its natural hardware capabilities.

In fact, this is exactly what has given us our edge over computers until now! Any useful formal system is necessarily very complicated and unwieldy. Hofstadter spends two chapters constructing a formal system for modeling number theory, and as a challenge asks the reader to write down the statement “x is a power of 10” as a string of the formal system. [Here's](#) the answer; scroll down to where it says “all together.” Humans have been able to outthink computers so far by implementing high-level heuristics instead of working with formal systems directly. And to the extent that computers are catching up, it seems like they're doing it *by coming up with even better high-level heuristics*, not by throwing more resources at more efficient formal systems. (Unrelated: upon reading an article saying that AI would put mathematicians out of business in 10 years, my advisor remarked: “That's just enough time for you to get tenure!”)

For now, it tentatively seems to me that Hofstadter is playing the “[Gödel's theorem and consciousness are both mysterious and therefore equivalent](#)” game. If there's some way to salvage Hofstadter's ideas in part II, someone other than me will have to write the book review doing it.



Relativity, by M.C. Escher. Like GEB, this drawing makes more sense if you only look at half of it.

Guns, Germs and Steel: The Fates of Human Societies by Jared Diamond

Guns, Germs and Steel by physiologist/ecologist/environmental historian Jared Diamond is an attempt to explain why technologically advanced societies developed faster in certain parts of the world. Its main thesis is that geography defines the constraints and opportunities for what a civilization can do: Groups of humans take their cues from their environment and follow these cues to develop tools and guns and spaceships or happily continue being stone age hunter-gatherers or anything else in between. Diamond sets out by tracing the different initial conditions that promoted the advent of agriculture and animal husbandry in different parts of the world and then follows the downstream effects and feedback loops that lead to Age of Discovery Europeans having the technological edge that allowed them to colonize the whole planet. In the meantime he presents accounts of the history of plant and animal domestication, the invention of writing, the biology of zoonotic diseases, the spread of technological innovation, the development of different modes of government, as well as crash courses in the history of New Guinea, China, Polynesia, the Americas, sub-Saharan Africa and Japan, and his ideas about the study of history in general.

The main part of the book begins around 11,000 BC, a time when everybody is a hunter-gatherer and nothing particularly interesting is happening anywhere. Humans have long spread through mainland Africa, Asia and Europe, have already reached Australia, New Guinea and inhospitable places like Siberia, have recently crossed the Bering Strait and discovered North America and will arrive in South America in another thousand years or so. At this stage (Diamond argues) it would not be obvious to a visiting Martian historian where technological civilizations are going to spring: Africa is the cradle of humanity and has the greatest human genetic diversity which could arguably lead to diverse cultures and inventions. The recently discovered Americas are huge and have great environmental diversity working in their favor. Eurasia is also huge, second only to Africa in its history of human habitation and already is showing some promise with its cool Cro-Magnon cave paintings. Australasians on the other hand have cave paintings at least as cool as the Cro-Magnons, have already invented boats sophisticated enough to have allowed them to reach Australia and New Guinea and have adapted to each successive island they hopped on. So what exactly was it that made Eurasia the right bet?

The first step down the road to technological civilization is to develop food production.

An agricultural society will have a denser population because an acre of farming will produce more calories than an acre of hunting and gathering. Population will rise faster as well, as nomadic hunter-gatherers have fewer children and longer spacing between births. Becoming settled farmers allows people to store food, so their societies have room for kings and bureaucrats and craftsmen and soldiers and other people not occupied with producing food full-time and free to focus on other things, like crafts and technology and cruelly subjugating their neighbors. For societies that domesticated animals as well as plants the effects synergize nicely: They can use manure as fertilizer for their crops and oxen to till soil that is too hard to till by hand.

Still, Diamond warns us not to make the stereotypical sharp distinction between nomadic hunter-gatherers and settled god-king-worshipping farmers, at least not straight away: No civilization suddenly woke up one day and decided to settle down and become farmers. It is better to think of agricultural practices not so much as a separate lifestyle but initially as a parallel evolved strategy for survival that was helpful in certain places and times. Many hunter-gatherers became settled long before adopting farming. Many cultures adopted some agricultural practices without abandoning hunter-gathering or switched back and forth between the two. Some hunter-gatherers actively managed their territory, for example by clearing weeds in order to protect the bushes they gathered from, in a way that set the scene for plant domestication later. Some early food producers planted seeds, went away to do some hunting and gathering for a few months and went back to the fields in time for harvest. The point is that adopting agriculture became *autocatalytic*; it led to positive feedback loops. As population rises, wild foods are gradually depleted. As more domesticable species are discovered, farming becomes more rewarding. Tools and techniques initially developed by hunter-gatherers are adopted and refined to improve farming. Gradually, groups that predominantly farm outcompete groups of hunter-gatherers, through the use of greater numbers and technology. As farmer civilizations expand, hunter-gatherers are pushed away or killed or forced to become farmers themselves.

The Fertile Crescent, a boomerang-shaped part of the Middle East covering the area roughly between the Levant and Mesopotamia, was the first place where the scale actually tipped against the hunter-gatherers. This was partly because that whole area was never particularly well suited for hunter-gathering in the first place but mainly because it happened to have the perfect mix of climate, topographical variation and presence of a vast array of domesticable plants (especially annuals) and animals ready to become domesticated: So by 6,000BC Mesopotamians were growing at least wheat, barley, lentils, peas and chickpeas, flax (to make linen) and keeping goats, sheep, pigs and cows; more than enough to feed a purely agricultural civilization, with wool, plowing and transport as added benefits on top of that.

Where else did agriculture spontaneously arise? Mesoamericans finally settled as farmers in a kind of similar way around 1,500BC, with maize as their primary grain (far harder than wheat to domesticate, so it took a longer time) and only dogs and turkeys available to them. At the same time the natives of the eastern U.S. domesticated the plants they had available, which turned out to be a type of squash, a type of spinach, sunflowers and sumpweed. Mexican corn finally reached them by 200 A.D., a version of corn that could actually be effectively grown in their climate showed up by 900 A.D., and two hundred after *that* (while Europe and the Middle East was going through the High Middle Ages) they got beans (also from Mexico), dumped sumpweed (on account of nobody knowing what that even was) and settled down to farm and form chiefdoms around the Mississippi. Other places fall somewhere between these extremes: China trailed the Fertile Crescent by maybe 500 to a thousand years with rice, soybeans and pigs as their most important domesticated species. The peoples of the Amazon river settled by 6,000BC but got agriculture around the same time as the Mesoamericans. Various parts of Africa like the Sahel region (just south of the Sahara), West Africa (between modern Ivory Coast and Nigeria) and Ethiopia also domesticated some of their local species (sorghum, african rice, palms, coffee) between 5,000 and 3,000BC, mostly after Fertile Crescent crops were introduced there.

(A quick side-note on animal domestication: It is no accident that only a handful of animals have ever been domesticated, even today. There are various reasons an otherwise domesticable-seeming animal species may be unsuited to being a farm animal, including their diet (for example carnivores), slow growth rate (for example elephants), unwillingness to breed under captivity (for example andean camels), panicky disposition (for example gazelles) or general tendency of the animal to be a complete asshole (for example zebras). The people of the Fertile Crescent lucked into having most domesticable species in their backyard. The people of Africa did not.)

Apart from the starting conditions, geography shaped the way food production spread: Looking at a map, on Eurasia most other civilizations will be to your east or west; on Africa and the Americas they will be on your north or south. This is important, as crops and livestock are adapted to a particular climate and climate differences are more dramatic along the north-south axis. Agriculture had an easier time spreading westward from Mesopotamia to Southern Europe than spreading southward from Egypt to the rest of Africa.

As Diamond describes it, once agriculture and population density reach a critical mass a whole lot of other things follow:

Germs: Living in close proximity with animals allows viruses to hop from animals to humans. You get epidemics of flu, smallpox, measles etc. but you also get populations

that over time develop resistance to the same viruses. Hunter-gatherers never got exposed to these germs and never got the antibodies for them, so an entire population of hunter-gatherers is in danger of being wiped out from an epidemic brought by a single outside visitor, as it happened with the native population of North America.

Technology: Settled, agricultural people have two big advantages in developing technology. They have the capacity to feed full time specialist craftsmen and they can accumulate more possessions than they can carry. Innovations breed more innovations, as people build upon or combine their technologies with new ones that spread to them from neighboring civilizations. Eurasia and North Africa got agricultural settlements early and, lacking any important geographical/ecological barriers for diffusing innovations, they developed faster. Western Europe took longer to develop, as innovations from the Middle East took some time to reach there. So why did Western Europe eventually surpass the Middle East? The answer lies again in geography: The Fertile Crescent was ecologically fragile and agriculture over time depleted its resources, turning it to mostly desert that we see today. Western Europe took initial crops and technology from outside to kickstart its food production but its land was more robust in the long term. So, as the Middle East and eastern Mediterranean slowly exhausted themselves the centers of power moved westward.

Writing: Writing is the odd one out, in the sense that for all we know it may have arisen independently only twice in history, by the Sumerians in Mesopotamia and by Mexican Indians in Mesoamerica. (China and Egypt are the other two possible independent origins, which may or may not have been influenced by the Sumerians). Many agricultural societies (like the Aztecs) never invented it at all. In any case it spread the same way all technologies spread, and although its use was limited at first it gave literate societies a huge edge down the line.

Institutions: Bigger denser societies inevitably develop centralized governments and institutions. After a certain population size, communal decision-making is no longer efficient and interpersonal conflicts can only be resolved through force monopolized by a central authority. Once civilizations get centralized governments, nearby populations are pressured to do the same, either through the threat of an external enemy or outright conquest.

Continue the above dynamics over a few thousand years after food production developed and you get the entire human history as a result.

So, what should we think about all of this? The explanation Diamond presents feels intuitive and compelling, at least in the broad, zoomed out view of history. It's also the kind of big picture explanation that feels particularly hard to think about and I am not

sure what is the right level of abstraction here. History is this super complicated experiment that ran only once and you have no way to replicate. So how important was blind chance in the whole scheme of things? How much of what happened is a pattern that is reducible to the geographical and ecological landscapes of 11,000BC and how much is noise? I am reminded of an anecdote about biologist J.B.S Holdane, who was asked by an overzealous Popperian student about what evidence might falsify evolution and he replied "rabbit fossils in the Precambrian". And looking at the various criticisms of the book, while I found several historians and anthropologists nitpicking or poking holes at Diamond's account of the history of various places (which I expected; if you summarize the 13 thousand years in 500 pages a few things will end up oversimplified or wrong), I didn't find anyone suggesting any conspicuous civilizations that developed in a way that directly contradicts Diamond's story. Criticism focuses on the book downplaying the importance of cultures or historical accidents, on it being too reductive, not on directly refuting its main thesis.

Or perhaps the entire book should be read backwards? Guns, Germs and Steel was written in 1997 and reading it more than 20 years later it felt that I already had absorbed some of its ideas by osmosis. Maybe it was more ground-breaking and surprising when it was written? The book begins with Diamond telling a story about how he was embarrassed by a New Guinea native named Yali who matter-of-factly asked Diamond how it happened that white people invented all the "cargo" (the local word for every material good brought to New Guinea by westerners, from steel and medicine to soft drinks and plastic toys) and his people invented so little of their own. This question brings up the whole terrible debate regarding possible innate genetic differences between populations, that this review will gingerly ignore. Diamond rejects it quickly, based mostly on his experience of having lived among New Guinea natives and finding them at least as smart as anyone else. The whole book is a reply to Yali's question. And Diamond presents it as the first convincing reply that does not require any assumptions about innate biological differences. And of course he has still been called a racist because he, to quote some critiques, "makes all the factors of European domination a product of a distant and accidental history" and "lets the West off the hook". So it goes.

And cultural differences must play a role, at least some of the time. There are examples of societies rejecting technologies for cultural reasons. Confucian China did not go about exploring the New World. Was its insularity a quirk of its culture? Or was the Age of Exploration a quirk of European cultures? Diamond makes an argument that China's geography predisposed it to political unity and the isolationist policy of its government applied to the whole region. Whereas the fierce competition of nations and exhaustion of Europe's resources encouraged a few of them to go about what eventually became the Age of Exploration in the West. In any case, this feels like a classic chicken and egg problem that is hard to pin down to anything more concrete

than “it was a mix of geography and culture kind of affecting each other in complicated ways”. But something like geographic reductionism must be true in many cases: Australia was populated by many different cultures, with completely different languages and lifestyles, adapted to every nook and cranny of the place; still, they were all stone age hunter-gatherers until Europeans showed up. Once the first humans stepped on Australia, unbeknownst to them, they locked all their descendants to the stone age as there was simply no ramp to anything else to be found anywhere in the whole continent. So, a lot of what we call culture is a downstream effect of more distant causes like geography or ecology, which may or may not then affect history in its own surprising ways.

I am not sure if I am convinced of every argument and every explanation of history that is presented in the book. But, kind of like the theory of evolution that I alluded to earlier, Diamond's theory makes a lot of different data points and intuitions click together in a particularly satisfying fashion. If we are to have a general framework to think about history in a global/general way, well I am not sure that Guns, Germs and Steel is necessarily the complete answer. But something pretty close to it must be true.

How Children Fail by John Holt

1.

Why are all children so bad at learning in school?

Seriously, they're terrible at it, and nobody ever calls them out as a group. We call out individual children as failing. We call out individual schools and school systems as bad. But the much more dramatic contrast is between learning in school and learning in any other context.

In their first five years, kids learn to understand 25,000 words, even if nobody is actively helping them, at the same time as they're learning most of what they'll ever know about physics, psychology, and how to pilot a human body. They then struggle to match this vocabulary acquisition rate over their next ten years, despite expert attention, a wealth of resources, personal encouragement, and even prizes.

After weeks of trying, my teacher gave up on getting me to correctly label the countries and capitals in a map of South America. Yet I quickly learned to navigate the New York City Subway.

Through involuntary cultural osmosis, I could probably pass a test on the characters, plot, and setting of *Twilight*, despite having never read any of the books or watched any of the movies. Yet there are books I read in school (good books, written to be enjoyed!) where I now couldn't tell you the main character's name.

(I was a straight-A student, by the way. Everybody fails at least this hard.)

As an adult, when I'm in the middle of researching something, and I get hungry, I get up, go to the kitchen, get a snack, come back, and keep researching. In most classrooms, children are absolutely forbidden from leaving to go get a snack because they would always claim to get hungry and they would never come back.

Which is not to say that children don't try, or don't care. They sweat and cry over tests. They care so deeply that they have nightmares about missing class that last well into adulthood. But by any standard other than comparing them to other schoolchildren, they universally fail. Why?

2.

In 1958, John Holt was trying to teach math, writing, and French to a class of ten-year-olds. It was going as this normally does--a large proportion were absolutely terrible at it. He wrote to a colleague,

If you live at a small school, seeing students in class, in the dorms, in their private lives, at their recreations, sports, and manual work, you can't escape the conclusion that some people are much smarter part of the time than they are at other times. Why? Why should a boy or girl, who under some circumstances is witty, observant, imaginative, analytical, in a word, *intelligent*, come into the class-room and, as if by magic, turn into a complete dolt?

He collected the memos and essays he wrote that year into a book, *How Children Fail*, meant to be read by educators and parents to help them understand what is going wrong in schools, as a first step toward figuring out ways of fixing it. Holt was a big believer in "hold off on proposing solutions until you've really analyzed the problem." When I first read this book, though, I wasn't in its intended audience. I was nine years old, which meant it was a handbook, not on education, but on rationality.

3.

More than thirty years after those students were frustrating Holt, I was the dolt in another small, private school. At home I was teaching myself to code in HyperTalk. At school I was exactly this kid from a Holt memo:

his school papers are as torn, smudged, ruffled, and illegible as any I have ever seen. The other day the class was cleaning out desks, and I was "helping" him. We got about a ream of loose papers out of the desk, and I asked him to put them in the notebook. As always, when he is under tension, his face began to get red. He squirmed and fidgeted, and began to mutter. "They won't fit, the notebook's the wrong size"-- which wasn't true. Finally he assembled a thick stack of papers and began to try to jam them onto one of the rings in his notebook, not noticing that the holes in the papers were at least a half-inch from the ring. As he pushed and fumbled and muttered, I felt my blood pressure rising until, exasperated almost to rage, I said loudly, "For heaven's sake, leave it alone, do it later, I can't stand to watch any more of it!"

My school held an annual fundraiser--parents donated used items, and the school sold them to other parents. There was always a pavilion full of unorganized boxes of old books. At the end of the day, students were allowed to take home any unsold books. We were told that any we didn't take would be destroyed, sent to a landfill. This was probably true (it viscerally horrifies me to this day), but if it was a lie it was a really clever tactic. It certainly motivated me to pick out books by the dozen that were out of my comfort zone--if they ended up unread on my shelf, that'd be no worse a fate than otherwise awaited them. It was a great engine of serendipity, and my reading this book was its best product--a book that no adult would have bought for me and that I probably wouldn't have ever checked out from a library.

Holt describes the bad habits of thought that schoolchildren fall into, many of which will be familiar to readers of 21st-century rationality texts. On positive bias,

Sometimes we try to track down a number with Twenty Questions.

...

They still cling stubbornly to the idea that the only good answer is a yes answer. This, of course, is the result of the miseducation in which "right answers" are the only ones that pay off. They have not learned how to learn from a mistake, or even that learning from mistakes is possible. If they say, "Is the number between 5,000 and 10,000?" and I say yes, they cheer; if I say no, they groan, even though they get exactly the same amount of information in either case. The more anxious ones will, over and over again, ask questions that have already been answered, just for the satisfaction of hearing a yes.

On failing to think concretely or across domains,

A friend was studying for a chemistry test. He was trying to memorize which of a list of salts were soluble in water. Going through the list, he said that calcium carbonate was soluble. I asked him to name some common materials made of calcium carbonate. He named limestone, granite, and marble. I asked, "Do you often see these things dissolving in the rain?" He had never thought of that. Between what he was studying for chemistry and the real world, the world of his senses and common sense, there was no connection.

And he makes similar claims, similarly argued, to those of [Paul Graham](#) and [Eliezer Yudkowsky](#), that the strategies that lead to nominal success in school are often the ones that stop at superficial understanding of the subject--hacks to be able to get to the

correct answer quickly, without ever really looking at the problem.

These lessons were helpful to me at the time. As someone generally much better than peers at getting the right answer on a test, I'd been eager to believe that that's what merit in school was. Holt taught me that I could get more out of class if I let myself be, nominally, a worse student: one who avoided memorization, who questioned what was being taught, who reached for generalizations beyond what would help with the next test. And for the areas where I already knew I was failing, Holt helped me a little to cope with the shame, anxiety, and avoidance spirals that grew out of and perpetuated that failure. Eventually I was able to use a 3-ring binder.

Per the last demographic survey of the readership of this blog, you are most likely not nine years old. However, you are almost certainly a *former* nine-year-old, and that's another excellent audience for this book. Holt, like Graham and Yudkowsky, sees school as instilling permanent cognitive biases--habits that are best unlearned whenever you can.

4.

Holt identifies three broad categories of failure modes.

Strategy - Perhaps uniquely for a rationality text, Holt identifies thinking strategically as a failure. Not in general--he encourages adults to ask themselves more often where they are trying to get, and then whether their current approach is getting them there. And he taught, often, through strategy games. But for a child, learning due to disinterested curiosity is much more effective than learning due to incentives. We're born with powerful drives to learn (learning could hardly start as a learned behavior). Replacing that with a desire to gain reward or approval, or to avoid punishment, or to even to Be A Good Student, distorts behavior. Children, consciously and unconsciously, start trying to maximize their perceived score. They seize on the most reliable way of figuring out the right answer, even if that's something overfitted to the classroom, like reading the teacher's face. They find a level of achievement they can reliably hit, then manage expectations to make sure they're never pushed to a higher one. If they're punished for giving wrong answers in class, they'll stop talking and disengage when they think they might not understand something. It's a familiar problem across domains--once you attach incentives to a convenient measurement of something, it stops being a good measurement. Readers of this blog may remember examples from machine learning or centrally planned economies. Holt's revised edition gives a quick one from the Navy:

I remember an old chief machinist on an obsolete training submarine in Key West

saying bitterly about his worn-out engines, which he had spent many hours polishing up for an official inspection, "They shine, don't they? Who the hell cares if they don't work?"

Fear - The more visibly dysfunctional response to incentives. Fear is the mindkiller, as another U.S. Navy veteran would write in a book published the year after Holt's. A student afraid of failure cannot acknowledge their own mistakes, and therefore cannot learn from them. Students labeled "Gifted" are so terrified of losing that label that they panic when they encounter something they're not excellent at, end up doing even worse, and try to avoid that whole area whenever possible. Holt describes a lot of his students as "emotionally incapable of checking their work"--not as a rational response to any incentive, but because looking for mistakes is like checking under the bed for monsters.

Boredom - Holt talks at length about children being bored in schools, but I'm going to skip over that for the same reason people rarely yawn in plays--you don't want it to spread to the audience. I don't think I need to sell you on the notion, anyway.

5.

All of these failure modes have the same cause. Learning in schools is externally imposed, not intrinsically motivated. School requires a power structure, with incentives and accountability, to cause children to reliably learn, to learn the Right Things, and to be able to prove that they did. In 1998, you didn't need to assign chapters to get one out of three kids to read *Harry Potter and the Chamber of Secrets*--they read it because they wanted to see what happened next. You didn't need to test them in order to find out how much they'd read and understood--they would insist on telling you about the book's plot, and ask questions about anything they didn't understand. You didn't need to find any way to get them to reliably sit still for it--they could be allowed to set their own schedule.

But say you wanted *every* kid who could to read and understand it. Then you'd need a school.

Similarly, you really don't need a school to get a decent percentage of children to learn math. I was as ravenous for extracurricular algebra as I was for extracurricular Harry Potter. But to get [Scott Alexander](#) to learn any not-directly-useful math, you seem to need a school. (If you replace "math" with "maps", our positions are reversed.)

The basic concept of school, then, requires incentives, fear, and boredom. And the fear component really isn't optional, says Holt, speaking from experience. In one of his

memos, he wrote

The other day I decided to talk to the other section about what happens when you don't understand what is going on. We had been chatting about something or other, and everyone seemed in a relaxed frame of mind, so I said, "You know, there's something I'm curious about, and I wonder if you'd tell me." They said, "What?" I said, "What do you think, what goes through your mind, when the teacher asks you a question and you don't know the answer?"

It was a bombshell. Instantly a paralyzed silence fell on the room. Everyone stared at me with what I have learned to recognize as a tense expression. For a long time there wasn't a sound. Finally Ben, who is bolder than most, broke the tension, and also answered my question, by saying in a loud voice, "Gulp!" He spoke for everyone. They all began to clamor, and all said the same thing, that when the teacher asked them a question and they didn't know the answer they were scared half to death. I was flabbergasted--to find this in a school which people think of as progressive; which does its best not to put pressure on little children; which does not give marks in the lower grades; which tries to keep children from feeling that they're in some kind of race.

I asked them why they felt gulpish. They said they were afraid of failing, afraid of being kept back, afraid of being called stupid, afraid of feeling themselves stupid. Stupid. Why is it such a deadly insult to these children, almost the worst thing they can think of to call each other? Where do they learn this?

Even in the kindest and gentlest of schools, children are afraid, many of them a great deal of the time, some of them almost all the time. This is a hard fact of life to deal with. What can we do about it?

Nothing, per the conclusion of the book:

The idea of painless, non-threatening coercion is an illusion. Fear is the inseparable companion of coercion, and its inescapable consequence. If you think it your duty to make children do what you want, whether they will or not, then it follows inexorably that you must make them afraid of what will happen to them if they don't do what you want. You can do this in the old-fashioned way, openly and avowedly, with the threat of harsh words, infringement of liberty, or physical punishment. Or you can do it in the modern way, subtly, smoothly,

quietly, by withholding the acceptance and approval which you and others have trained the children to depend on; or by making them feel that some retribution awaits them in the future, too vague to imagine but too implacable to escape. You can, as many skilled teachers do, learn to tap with a word, a gesture, a look, even a smile, the great reservoir of fear, shame, and guilt that today's children carry around inside them. Or you can simply let your own fears about what will happen to you if the children don't do what you want, reach out and infect them. Thus the children will feel more and more that life is full of dangers from which only the goodwill of adults like you can protect them, and that this goodwill is perishable and must be earned anew each day.

6.

Holt never stopped looking for ways to make school incrementally better, but the basic tragedy seems baked into the very concept. We have things we want all children to learn, so we need institutions to teach them. Children on their own are curious, but not driven to learn exactly those things, so we need ways to encourage them. Institutions need accountability, so we need ways to prove that the children are learning those things. So we impose incentives and legible performance metrics, and that kind of breaks kids. And we tend to stay broken for the rest of our lives.

Holt wasn't always explicit about this, at least until the revised edition. He was still employed by schools, still trying to make schools work, and therefore reluctant to characterize schools as outright destructive. But not *that* reluctant, and the connections are inescapable. Failing students blindly follow "recipes" to solve math problems, too insecure to stop and think about why the recipes exist or whether they're moving in the right direction. Failing teachers use traditional pedagogy even when it's demonstrably not working, too insecure to even consider they might be making things worse. Fear and insecurity beget fear and insecurity down the generations, with bullied students growing up to become bullying teachers. Students conditioned to think of schoolwork as an abstract character-building exercise become parents who don't care whether what's taught in school has any meaning or utility. I suspect this is at least a little true and is why it took a frustratingly long time for me to persuade my parents to let me drop out of school. (I ended up "self-schooling" for three years, from 15 to 18, before going to a traditional four-year liberal arts college. It was great.)

7.

The long-term psychological effects of school are, as far as I know, ambiguous. Perhaps it damages our rationality or disinterested love of learning. Perhaps that's just a

natural part of growing up, and schools help us develop the skills that we'll need to do work we're not always interested in. The child welfare case against directed learning is also hardly a slam dunk. Plenty of children like the control structure in school; it's gamifying something that's already fun. But if we set those issues aside for a moment and zoom out, there seems to be a clear tradeoff between efficiency and control. Children are naturally learning at high velocity, but may not be learning the most useful or important things they could be. You can exert control to optimize the set of topics they're learning, but at a cost of reduced velocity.

Holt gives an anecdote of a fifth grader caught sneakily reading a science book when he was supposed to be learning about "Romans in Britain." By forcing him to put the book away, the teacher traded an hour of high quality science education for an hour of low-quality history education, during which the child is less engaged and will remember less. But perhaps the marginal value of the latter was still higher; maybe knowing one thing about Romans in Britain, rather than zero, will be more valuable than knowing a hundred more things about a scientific topic you've already gone deep on. When you take away the babysitting and socialization benefits of school (as has happened recently) this marginal value proposition is what's left. You might not buy it for the Science vs. Romans story, but surely forcing kids to learn just a little reading, writing, and arithmetic is worth it in the long run, no matter what else they'd rather be learning. But we've taken this to an absurd extreme, where self-directed learning becomes a tiny part of a typical childhood, something you carve out a little time for during your nth year in state-mandated courses of study you have no interest in. Why did we all have to learn that the mitochondria is the powerhouse of the cell? That's not even the most interesting or useful fact *about mitochondria*. Society does not depend on universal lifelong literacy in cell biology; we could get along just fine if 5% of us knew the powerhouse thing, and the other 95% had skipped bio class that year. I assert we would, in fact, be better off, because those 95% would have learned more. Our desire for scalable, testable instruction leads, by middle school, not to ensuring breadth of learning, but to forcing identical amounts of depth in the same set of somewhat arbitrary subjects.

8.

How Children Fail was published sixty years ago, when the most exciting piece of educational technology was [a set of wooden blocks](#). John Holt wrote a revised edition of How Children Fail in 1982, after becoming an advocate for homeschooling and child emancipation. The original edition ended by calling for reform of the school system, saying it should provide buffet tables for students to browse, rather than force feeding. The revised edition adds a postscript:

[E]xcept in very rare circumstances the idea of special learning places where nothing but learning happens no longer seems to me to make any sense at all. The proper place and best place for children to learn whatever they need or want to know is the place where until very recently almost all children learned it--in the world itself, in the mainstream of adult life. If we put in every community, as we should (perhaps in former school buildings), resource and activity centers, citizens' clubs, full of spaces for many kinds of things to happen-- libraries, music rooms, theaters, sports facilities, workshops, meeting rooms- -these should be open to and used by young and old together. We made a terrible mistake when (with the best of intentions) we separated children from adults and learning from the rest of life, and one of our most urgent tasks is to take down the barriers we have put between them and let them come back together.

He also notes in passing that computers (“though *only* if very different from present ones”) might be able to help children clarify their thinking. Well. Tragically and frustratingly, Holt died only three years later, and so he didn't quite get to see computers become that mixed-age community center he'd envisioned. I was a privileged kid, and when self-schooling I had access to all sorts of resources. But all I really needed, back in 2000, was AOL. Technology provides education in attractive snack form and allows special interests to scale.

Screens are an incredibly strong draw to anyone over, say, four or five years old. You can get a four year old passionately interested in learning to read by making literacy tools the only video game available. I don't know if Holt would fully approve, but you don't really need a control structure anymore to get kids to study the exact thing you want them to--you just need to make the only sophisticated electronic device readily available be one programmed to teach it. Then, a little later, introduce them to the concept of edutainment videos on YouTube, and for better or for worse, The YouTube Algorithm will take it from there.

The internet's collapse of scale also removes much of the need for standardized education. Holt talks about a fifth-grader who was fascinated by snakes, wanted to know everything about them, and learned more math when it happened to be incidentally involved in snake facts than in a math class. Snake Kid's school couldn't be expected to create a whole Snake Curriculum just for this one student, so if Snake Kid wanted an adult to explain things to him, those things would not be Snake Facts. But today, an endless series of snake videos is just a click away.

And the internet is a far more powerful engine of serendipity than a pavilion of unwanted books.

9.

School is an institution. We come to depend on it, it defends itself, it's not easy to change from inside or out. I mostly believe the critiques in Holt's book, but I generally haven't felt that radical reforms were possible. I've encouraged children who don't need as much socialization or supervision to consider getting parental support for dropping out of school, and that's been more or less it, reform-wise. But we maybe do have a moment here, in 2021. We've had a massive disruption in traditional schooling, which could be an opening to escape from inertia. We'll at least have a natural experiment as different schools physically reopen at different times.

What would happen if we let kids choose how to allocate more of their time, and gave them the support and resources we could? What if the free daycare provided by state schools wasn't as coupled to regimented instruction, but still included books, computers, and adults who would help explain things if asked? A lot of kids would spend their time in ways that didn't look very productive to us--binge watching TV, inventing new sports, making a thousand paper cranes. Almost all of them would end up with educational deficiencies that looked shocking to us. I am skeptical that in this day and age the consequences would be dire, though. It's really easy to catch up when you need to (did you know you can just *look up* what the powerhouse of a cell is?). With less Educational Attainment gatekeeping down the road, we could more safely test this. And I really think we should, because while the benefits of schooling are ambiguous, some of the costs are not. Kids are definitely less free to pursue their passions. Unless you have a really exceptional amount of energy and parental support, school puts a ceiling on your personal productivity as a child. And long-term consequences aside, school is just not maximally fun. Your greatest capacity for joy was as a child. The time of a ten-year-old is just as valuable to her as the time of an adult is to the adult. And yet, even when a classroom format isn't working well for a ten-year-old and they need some kind of other instruction, we often still stash them in a classroom to kill time being miserable and frustrated in between actual study sessions, as if there's no benefit in them having fun.

Suppose you were asked to spend a day as a fifth-grade student. Six hours, with breaks, getting taught things you mostly already knew, with an adult in the room who will try to catch you not paying attention and at minimum be offended if she does. You might do it for the novelty or nostalgia, I suppose. For a day. How much would someone have to pay you to do it for a week? For a month? How good a case for the social utility of the thing would somebody have to make for you to do it for nine months, for free? I was a lower-case-g-gifted fifth grader in a school without a Gifted program, and when I complained that my time was being wasted this way, nobody cared as much as I

thought they should, because nobody thought there was anything else worthwhile I could be doing. I'm glad I had this book back then to tell me they were wrong.

Human Scale Revisited by Kirkpatrick Sale

I

I read Kirkpatrick Sale's book *Human Scale Revisited* with some trepidation. Pop culture has conditioned me to side with the plucky rebels in most fights. They're the plucky heroes who fight against the empire. They're the heroes when fighting an illegal alien occupation. They're the heroes even when they're green or blue or turquoise in a planet far far away, which nobody can remember because it's slunk away from the cultural consciousness.

So when I picked up Sale's book about the joys of simple living and the curse of Bigness, and its pocket companion *The Collapse of 2020*, I wondered if I would end up going full Thoreau and throw out my computer and phone. I disagree with almost everything that a neo-Luddite believes, but the temptation is a strong one.

The circumstances surrounding the author and the book were arguably more interesting than the book itself. Sale had made a bet with Kelly, of *Wired* fame, about whether we'll face a civilizational collapse 25 years hence in 2020. He bet that three things would happen - the global currency collapse, significant warfare between rich and poor, and environmental disasters of significant size.

At first when I had read of the bet I dismissed it out of hand. Who wouldn't? I had lived in 2020, and it didn't seem like I was in the middle of a civilizational collapse. I'm sure I would've noticed.

But it's still helpful to see how the luddite beliefs of *Rebels Against the Future* has evolved over a quarter century to a short and rather depressive conclusion.

It's not just the relentless pessimism that makes the book interesting. It's the fact that half of the extrapolations sound farfetched, but half of them sound *exactly* like the things that smart, educated people worry about today.

For instance, take South Bronx in New York City. It started its life as Morrisania, the private domain of the aristocratic Morris family, including Lewis Morris who was a signer of the Declaration of Independence. As they developed their landholdings, German and Jewish immigrants came to the area. It was called the Jewish Borough, with almost half of its residents being Jewish in the 1930s.

Then came World War II, and post that the feverish work from the federal government, which changed the entire face of the neighbourhood. Literally. It went from two thirds white to its opposite in about a decade from 1950 to 1960. There was landlord abandonment and white flight, coinciding with economic changes, crime and all supposedly influenced by the construction of the Cross Bronx Expressway by Robert Moses.

It cut through the heart of Bronx and displaced thousands of residents. Some neighbourhoods were entirely destroyed. This meant that those who could leave, did leave, and created a bit of a negative spiral. The Civil Rights movement and the general racial tension didn't help matters much.

Property values naturally plummeted, which brought more flight, which reduced it further, which reduced economic opportunities, which brought in more crime, which reduced the property values. And so it went.

All this, Sale argues, helped create what's effectively a modern urban slum. As states grow, they swallow cities and villages. And the consolidation of power at the state level means there's little power left in local centres. This starts off the downward spiral, of trust erosion, of a lack of self-sufficiency and cooperation.

Never before have nations grown so large, never have corporations become so powerful, never have banks and brokerages been so influential and intrusive, never before have governments swollen to such unmanageable sizes, never have the systems, the factories, the technologies been so huge—hence, never before have the crises been so acute.

As a sort of Dunbar number for cities, Sale suggests a maximum size of 100,000 inhabitants. Ideally it should be smaller communities of around 5000, with your own "commune" equivalents of around 500, which can be seen in ape colonies also as the ideal maximum size for a group.

The core of most of the problems is that we have a government that is much too large. And he continues in fine invective, laying out his wrath against the seemingly expansive government, which also lays out his worldview in a nutshell.

Governments, whether meaning to or not, always seem to create more havoc as they grow larger, and the largest of them historically have tended to be the most destructive and bellicose. ... Indeed, so regularly does one encounter this phenomenon in the reading of history that I am emboldened to advance this as a full-blown maxim, what we may call the Law of Government Size: Economic and social misery increases in direct proportion to the size and power of the central government of a nation or state.

Why is this growth malign, in Sale's view? He relies on Lewis Mumford, the historian of cities, to help him out here.

Throughout history, he has shown, the consolidation of nations and the rise of governments have gone hand in hand with the development of slavery, the creation of empires, the division of citizens into classes, the recurrence of civil protests and disorders, the erection of useless monuments, the despoliation of the land, and the waging of larger and ever-larger wars.

Anyway, what other kind of problems, apart from the odd insane Expressway, do governments cause? Sale has a list. It's long.

For instance, when we used to have medieval guilds (or if you're in London, as you still have), the guilds had an oath of assistance towards each other. A sort of multilateral treaty that they'd take care of each other. But when the organisation was destroyed by the King in the interests of mercantilism, it destroyed that network. And what happened? The smaller craftsmen and workmen fell through the cracks. And once they fell through the cracks, the government stepped in with charity and offers of help, building and operating hospitals and a welfare state, slowly but surely increasing our reliance on the

government. Thus goes the dystopian tale of "how the government made us dependent on them".

And if economic ruin and a creating a malign co-dependence wasn't enough, what else does the Big Government get up to? They wage wars.!

In fact, Sale argues that governments are much more likely to wage large wars when they're unified and large, rather than when they are tiny little duchies. There are sections that detail the multiple large scale conflicts that exists in the modern world (in 2015, 65 countries fighting wars and 638 conflicts between insurgents and separatists), which didn't happen in our bite-sized history.

It is an interesting fact that when the peoples of Germany were divided into dozens of little principalities and duchies and kingdoms and sovereign cities — from about the twelfth century to the nineteenth — they engaged in fewer wars than any other peoples of Europe. ... Not that there was total peace, nothing so otherworldly as that. But there were long stretches without war, and those (mostly internecine) wars that did erupt tended not to be so intense or so lasting as those on the rest of the continent. All that changed, of course, with the unification of Germany and the establishment of one government over 25 million people and 70,000 square miles."

(I think Sale is confusing the frequency of wars with the cumulative damage of wars, but point taken!)

He also is concerned about a steady increase in dictatorships and authoritarian regimes, on government corruption and cronyism and wastage of extreme proportions.

And, that's not all. The fact that the governments are larger today means that they're actively the creator of several new social problems. They're more inefficient, autocratic, wasteful, corrupt and harmful. The idea being that this impersonal apparatus, like in Kafka's dream, causes irreparable harm on smaller polities of all shapes and sizes. Sale quotes an economist from Warne Crews of the Competitive Enterprise Institute (which sounds wholly unbiased though I haven't been able to track down the exact calculations) that the government regulations cost the economy \$1.89 Trillion in a year!

This growth of government that then gives rise to a multitude of problems gets the catchy name of *prytaneogenesis* - where the seat of government creates an illness, akin to an illness caused by a doctor.

And if that's not enough, in a sane world:

It would not allow the production of 89 million polluting motor vehicles (2014) every year, or countrywide fracking that fouls drinking water and creates earthquakes, or metropolitan areas of 24 million people, or a cabinet department (Homeland Security) formed out of 22 agencies with 216,000 bureaucrats, or the manufacture of 387 cereal brands in America, or a Code of Federal Regulations that at 175,496 pages in 2014 was 117 times as big as the Bible, or a single World Trade Organization, governed by a secret court, regulating 90 percent of international trading.

Where else does the government's grubby hands cause disaster? It's in our corporations of course. Sale's argument is that in this mechanistic age that was kicked off by the Industrial Revolution we are increasingly getting detached from our humanity and actively screwing up the planet. His actual list of complaints is a little larger:

There is as well an imperiled social ecology, with a breakdown of family and community throughout the industrial world, an increasing dependence on the uncontrolled internet and "social" media, international cyberwarfare, unchecked and spreading slave and sex trades worldwide; in the United States, an erosion of religious commitment, contempt for law and law enforcement, increasing alienation and distrust of established institutions (throughout the end of the twentieth century and for all of the twenty-first, the General Social Survey of the University of Chicago determined that public confidence in all American institutions is below 30 percent), cultural ignorance and confusion, ethical and moral deterioration, a growth of suicide, mental illness, alcoholism, prescription and nonprescription drug addiction, alienation, poverty, domestic and racial violence, broken and divorced families, and unwed motherhood (from 4 percent to 40 since 1950, 93 percent of mothers under 20 in 2015), an inability to establish racial harmony or justice, and increasing rates of mass incarceration by an order of four since the 1970s.

This isn't just an isolated screed. Sale is arguing for a new world, one that's more tractable and less Tower of Babel compared to ours. Surely companies have economies of scale and marginal cost reductions at least, so just in this one teeny tiny instance bigger would be better right?

Wrong.

..in fact, as Barry A Stein, the leading scale economist, has shown, "studies of productivity, as measured by value of shipments per manufacturing employee, show that the highest efficiencies tend to occur not in the largest size plant, but in those of moderate size," that is, *with fewer than forty-five people*.

And he's not content with destroying the core of economics by killing the idea of economies of scale. Our love of bigness doesn't even extend to farming, where smaller farms yield more at a higher value and also enrich the locality. Where small farms have higher yields than large farms, and in any state the value of the crops grown on the average acre tends to be larger when the average farm is small.

And what's Sale's solution to these giant corporations and their hegemony? Well, it's having smaller companies. And have it be worker owned. It has this example which hit home. In eastern Pennsylvania, around 1973, one section of the mine established an autonomous work group. The workers were well trained, they'd set their own hours and they didn't have any bosses to tell them what to do.

And not only did the production quotas not decrease, accidents fell 5x, safety violations fell by 50%, and the miners were actually happy. Yup, somehow they managed to make happy miners an actual thing!

But along with the government causing problems of size and corporations also causing similar problems, we're not done. Did you know that 15k scientists from every country

around the globe got together in 2017 and warned all of humanity that we're doomed? Or that there's only 5% of wild animals are left globally.

For instance, there are sections dedicated to the economic collapse. In an explanatory article, he lays out a slightly condensed thesis.

The ten hottest years on earth have been between 2005 and 2020, with 2019 the hottest ever recorded and 2020 very close. That means ice melting at a record rate, with significant loss at glaciers around the world, in Greenland, and at the poles, with ice going three times as fast in the last three years in the Antarctic as just ten years ago and the Arctic in what a scientist at the Polar Ocean Physics Group at Cambridge University has called a "death spiral." The U.N. climate panel, which puts the blame for global warming on "greenhouse gasses," says these must cease by 2030, a goal that not a single major country is capable of meeting. Add to this the assault on the world's oceans through acidification and overheating, including 60 per cent of the world's fisheries fished to capacity and 33 per cent overfished, and the extinction of species at a rate that one scientific team in 2017 said offers "a dismal picture of the future of life," and it may fairly be said that an ecological collapse is well underway if not yet quite complete.

But that's not all. The larger governmental institutions we've set up, like the WHO, brought DDT and the plague to Malaysia as a side effect of their zeal to eradicate malaria. This is where Sale reaches inwards and pulls out his own James Scott, and starts sounding eerily similar to *Seeing Like A State*. For instance, he goes through, in powerful detail, an example of peanut farming by the British ruining thousands of acres of land in Eastern Africa.

Sale explores how the income differences are vast and getting larger, how the middle class is shrinking, how there is ineradicable poverty around the globe and how there is an incredible amount of consumerism that's only making the richest richer.

Ok. we've gone a long way to establishing that we've dug ourselves into a nice hole, what's the solution? Sale has the answer.

So to save our planet and its civilizations we must move in an opposite direction, we must work toward the decentralization of institutions, the devolution of power, and the dismantling of all large-scale systems that have created or perpetuated the current crises. In their place, smaller more controllable, more efficient, more sensitive, people-sized units, rooted in local environments and guided by local citizens. That is the human-scale alternative.

Technology is like discovering magic in Middle Earth. The world as you know it changes once it's here. While it's calming to think of the (imaginary) idyllic days gone past, there's no reversal of the arrow of time.

The key issues he sees span all worlds - the economic order that is ripe for collapse, the political order is deteriorating in front of our eyes, and the environmental collapse which is becoming large scale and inevitable.

He also talks about our life being fundamentally different, in scale and scope, to what came before. Our cities are vast, our crises are different and nothing in the world can stem the tide of exponential growth. To Sale, exponential growth is cancer.

To put it simply, modern technology enables us to do the bad things we're doing faster and more efficiently than ever before.

I am of course deeply troubled that those I know and love may not continue to exist, that many benevolent and righteous souls may vanish, and that much that was good in human achievement and art is gone forever. That is nothing but sad, yea heartbreaking, for any sentient human. But when I reflect upon what that species has done, what fundamental evil it has thrust upon the world to cause its civilization to collapse, I cannot but feel that it is ultimately for the betterment of Gaea that it should cease to thrust its pernicious presence upon the earth.

And naturally since environmental, political and economic catastrophe comes from our curse of Bigness, where the government and corporations have become all powerful, the only solution is its removal.

What's needed is massive decentralisation. Much like the kings of yore who put down their swords and armies and walked away, he's asking humanity to do the same. To get to smaller, more efficient, more controllable units of society where people can have rooted lives.

After all, if you want to survive stably for a long while, you gotta go small. It holds true for countries (San Marino, tiny and thriving since 301 AD), companies (Keiunkan, a Japanese hotel operating since 705 AD), organisations (I couldn't figure out if the Catholic Church counted) - the trend is clear. Athens, not Rome, that's the motto.

Sale also has a small aside where he explains why the English system is preferable to the metric system - because the English system of an inch (distance of the first joint of the thumb), a foot (length of forearm) or the yard (extended arm from fingers to nose) is a human one. The metric system relies on an impersonal calculation, not human at all. This despite the fact that not all of us have the same thumbs or forearms. It's not a major part of the book, but amongst all his recommendations this is the one that's likely to stay the longest with me. And now you.

II

To start with, there are a few major pros that you have to give the book - it's relentlessly optimistic (in a completely pessimistic tome otherwise mind you) that humans can live happily and sociably and lovingly in small communities. To me that vision is somewhat of a comfort blanket, because it's something I'd like to believe too. I wanted to be one of the Na'vi, like everyone else, happy and blue in my paradisiacal land.

But still. There are some clear sins that Sale commits in his books. Let's have a look - they fall into mostly four buckets - 1) his fear of exponential growth, 2) his misunderstandings of how networks actually form, 3) an inability to separate the worst of the modern world from the best of the modern world, and 4) a misunderstanding of where innovation and progress actually comes from.

So the **first misunderstanding**, his fear of exponentials. This is perplexing because instead it should be his friend. For example, there are stretches that wax lyrical about the Golden Age of Athens when we constructed the magnificent beauty of Parthenon. But would you really skip back 2500 years to give there? With no healthcare or Netflix or halfway decent epic poetry? When you read about Socrates taking hemlock it's a grand

gesture. When you're standing in a smelly street corner wishing they had invented coffee, life's a tad different.

Everything bad he denotes has indeed come with exponentials. But so has everything good. For things that build on other things, which include population growth but also includes science and technology, exponentials are great. Even with things like population growth, which seemed uncontrollable and scary when I was growing up, the world seems far more capable of supporting everyone, with the more brains the merrier. And if he wants to throw out the Romer-style endogenous growth argument with the bathwater, that seems a tad harsh.

So if you did want to live an idyllic life today, you should in fact embrace everything that technology can give you, and look towards shaping the world as we know it. Our cities are different not just because of size, but because we now can have different cities than they could in the past.

This is the **second misunderstanding**. Even while the physical city might have a few million inhabitants, they're all inhabiting a virtual world filled with friends and family and colleagues from dozens of different locations. And that's a scale advantage. It's one that Geoffrey West talks about a lot in his research, how both good things (ideas, startups, wealth) and bad things (robberies, crashes, traffic jams) happen way more in cities than otherwise. I'd argue the answer is not to run away from the super-scaling that's going on but to understand it and make it your friend. That's what we've been trying to do anyway.

Also, Athens in its glory days wasn't small by choice. It was small by circumstance. Choosing not to grow is not a luxury our species has ever shown much proclivity for. In fact, the only real modern society that has embraced this contradiction might be Japan, but even there you can see robot battles and skyscrapers galore!

The **third misunderstanding** is more a bias issue. Sale focuses so much on all the ways in which our worship of the "large" has caused issues that he forgets that it's not a solely bad thing. To his credit he recognises the bias, but once recognised he puts it behind him, forgets about it and continues on anyway.

Would you rather live in a smaller town? Well, it depends. Do I still get cafes and bookshops and restaurants and movie theaters and takeout at 1am? Do I have the chance to bump into complete strangers who might change my life? As much as all TV series show idyllic lives in small town settings, they brush past the reasons why the main characters leave for cities in the end. It's for opportunity. Humans like to congregate where other humans are, because that's where the action is.

Because for Sale it's a fair bargain to give up our trappings of modernity in exchange for an idyllic small town life. And you know what, that's totally fair. The problem is that this isn't a one-man equilibrium yet. Most people won't. Or don't. Or can't.

If we did want to live in "human scale", we could do that through going back into Athens, when you could know everyone, or you could go to modern day Silicon Valley, where much the same ethos lives. The difference is that only one of those places has indoor plumbing.

And the **fourth misunderstanding** is an existential issue. There are reams of work trying to uncover the secrets of innovation and progress. There are prominent folks asking more people to take on the mantle and try to look at the problem. My small efforts included, this is an existential problem that humanity has to master to get to a Star Trek era, where we can have the life that Sale dreams of, just with matter replicators.

But to get there requires that we, as humanity, work together. We need [genius clusters](#). We need better progress narratives that [detail our technological ladders](#). We need more scientific funding, increased experimentation, more ideas, and better cross functional collaboration. That's unlikely to happen if we're not also living in a society where "Bigness" was helpful.

It's a collective action problem. The US spends \$130Bn in Federal R&D a year. That amount can be spent because it's decided by a centralised entity. Split that out amongst a 1000 entities, and you have organisational and coordination chaos.

III

Ok, so he does make some mistakes. But that's fair enough. But as he's a neo-Luddite, he's probably also the best person to see the flaws in *our existing* system. And that he does with aplomb.

He explores some ideas on decentralisation that seems straight out of the Peter Thiel playbook. Cities should have more autonomy and funding. It needs to be directly responsive to its residents. (He would probably approve of the Miami Mayor and his direct Twitter game.)

This is very much in the Jeffersonian plan for smaller republics that came into vogue in the early 1800s. He called it "ward governments", small units of a few hundreds to a thousand, as separate units. It's also echoed by Thomas Paine, who remarked how there was a Pax Americana for a long few years in multiple US states post the American war when there also was no government. To him, this was no coincidence.

He also tries to look at the various city states, including Singapore, Monaco and The Vatican, to see if they resemble Ancient Greece, or at least Athens, in their cooperative spirit and strong network of communities. Turns out, kind of.

This part of the defense also sounds remarkably like the writing of Mises. A long litany of the inefficiencies of larger governments, who are essentially large, blind, cantankerous, wasteful entities, and the list of ways in which their ills are exacerbated by their size.

What Sale does in the book is to make a pretty strong case of a world where we could be living. One that reduces the sense of alienation we feel and increase the sense of connection.

It's a utilitarian calculus. If we're harming the world this much, and ourselves too, then perhaps the progress we're getting is not worth it. In Tyler Cowen's *Stubborn Attachments* he talks about the optimal belief in "sustained economic growth" as what humanity's north star should be. This would be a case where he will even get a

technological pessimist to agree. Just that they will have slightly different definitions of the world "sustained".

The ultimate goal that Sale has, as is evident from the title, is to push the theory that we should live in a world that is comprehensible to humans. Buildings at human scale, societies at human scale and whole economies at human scale. You can almost see his frustration at me reading this impersonally on a Kindle.

One of the benefits he touts, accurately, is that a sense of belonging is great for humans. It's why religion is good for us. The book is a paean to tell us all to pitch in to society!

Which conveniently ignores the fact that we do! Volunteerism is [pretty high](#)! 30% of adults volunteered through an organisation in 2019. That's not including all the people who are active in communities, join thousands of clubs which now, thanks to technology, they can.

How will tiny city states defend themselves against avaricious states?
Avaricious states fight far more than tiny city states, so if everyone's in a tiny city state then there would be less wars, period.

Okay, but that doesn't feel like much of a Nash equilibrium does it? After all, the age of city states in days past also saw Genghis Khan and Alexander.

Yes, but those are isolated examples from a 2000 year old history. Instead the small duchies can just band together with a mutual defense treaty and create a deterrent.

Right. What about the fact that you need large states to deal with pollution or healthcare?

We don't need the federal govt to take those on! We can deal with that perfectly at the city state level. In fact, because of *metis* it'll be faster and better.

IV

All that said, I do have a soft spot for both the books. Playing Cassandra is a hard job. Playing Cassandra when half your predictions come true is even harder!

And in many ways I do believe in Sale's vision. Of smaller, livable communities, of parks and woods and animals, of the feeling of communal belonging. I'd like that.

Not to mention the fact that his annoyance that the world is no longer easily tractable is very widely shared! Our world doesn't just scoff at wannabe polymaths, it ensures that confusion around life is a core condition. We are so steeped in complex interrelationships that this is not a choice we're given.

However, there's also a false dichotomy at the heart of the narrative. Some of the strongest technologists of the time argue for more decentralisation and community based living.

As technology evolves, the most profound and destabilizing change is likely to be the transition from centralized internet services to decentralized ones.

[India through crypto] defends national security by preventing deplatforming, deters fraud via on-chain accounting, and offers a decentralized alternative to a new Cold War.

While the crypto-speak is the current zeitgeist, the feelings have been a long time coming. You might recall Peter Thiel tried to do seasteading to set up a marine nation once upon a time not that long ago.

The natural endpoint for libertarians is to voice support for increased decentralisation and more localised governance.

Yet, in any instance, technological rise and economic growth do have externalities. In fact, wrangling those externalities is kind of what a large part of the liberal order is working on now. Granted they're not particularly fast, but that's the battle that's worth fighting.

The funny thing is that this book, so clearly dystopian, was about a hair's breadth away from being utopian. With minor changes (perhaps the title) we could've had a book that talked about where we wanted to get to, and talked about how the arc of economic history has made living in smaller communities arguably *more* possible today.

The fact that life seemed more idyllic in Athens circa 500 BC tells us about what we're missing, and once you finish snorting, about what the next wave of decentralisation is likely to bring us.

More interestingly, the question that it raises is an important one. What is the right type of ecosystem we should aim to design for ourselves as humans?

This has profound implications in how we think about our world. Paraphrasing Douglas Adams, using feng shui to create a space that makes a dragon happy might also make us, another organic lifeform, happy, and the fact that dragons don't exist doesn't matter one whit to the system being devised.

Similarly, the fact that we're unhappy in several ways as detailed above might result from our, dare I say, alienation from the modern world. One of the implications of the fact that the world has gotten larger is that our tools to wrangle it to become smaller haven't worked.

And this problem that was created by technology can only be solved by technology. After the world got big when everyone schlepped off to four corners, we discovered magic flying machines to help travel faster. And then we discovered even more magical talking rectangles that almost eliminates distance altogether, at least insofar as connection is concerned.

We do live in digital villages today. The fact that they're not physically contiguous means that not all of our needs gets taken care of in the same fashion that Sale imagines. It's not the same as living in a village in the Cotswolds. But it's still pretty amazing!

One of the problems with the modern zeitgeist is that it's not easy to read dissenting voices at the base level of the system. We don't have enough Marxists tearing at the fabric of our societal beliefs, and keeping us honest. Which is why reading Sale's work was invigorating. It left me far more hopeful about our future than it had any right to do. I think you should share in the enjoyment.

Human Smoke by Nicholson Baker

History is not a morality play, even though on occasion it is taught as one. *Human Smoke* by Nicholson Baker isn't so much a traditional book as a collage of footnotes moving through time. Set in the 1930s and early 1940s, a typical page of the book looks like the following:

Churchill visited Birmingham, where eight hundred people had been killed in a raid a few days earlier. It was the end of November 1940. *"A very pretty young girl ran up to the car and threw a box of cigars into it,"* Churchill later wrote, *"I was very glad (in my official capacity) to give her a kiss. I then went on to see the long mass grave in which so many citizens and their children had been newly buried."*

That's it. A line break and a new little story unfolding somewhere else in the world. Baker looks to fill in all the details that get wrung out of our current historical telling of World War II, many of which do not reflect well on the Allies. In this way, the book is a bit like Howard Zinn's *A People's History of the United States*, except much better. Better in the sense that the true story is shown - not given - and better in that the ideological shading is much harder to see.

Besides the structure, the other artistic choice is the scope Baker chose, both in terms of time and subject. The book focuses on the time period just after World War I and ends just after the entry of America into World War II, cutting the action of the war in half. You never see the liberation of the concentration camps, the use of the atom bomb, or D-Day. Its subjects are also somewhat oddly chosen, with emphasis on Churchill, Nazi High Command, European Jews, President Roosevelt, pacifists, and on the tactic of mass aerial bombing. Baker conspicuously underplays the Soviet Union, the use of tanks/blitzkrieg, the fall of France, the Japanese invasion of China, and Mussolini's Italy. While these events are mentioned, they are only mentioned briefly and without much depth.

The limited scope is what ultimately makes this book a piece of art rather than pure history. Baker has an argument, but he isn't going to come out and tell you what it is. *Human Smoke* is not a total retelling of World War II and is not meant to be complete. Read between the lines, however, and what emerges is less the valiant heroes taught in civics class and more the very real and very fallible people who had to make decisions in times of war.

Aerial Bombing

More than anything else, this is a book about the love of bombing. Throughout every major event and decision, a bombing raid was always looming in the background. If it wasn't a bombing raid, it was production numbers on bombers, or a look at the type of bombs that were dropped, or what new technology had developed to make bombers even deadlier. Questions

were debated like: What was the right mix of incendiary versus traditional bombing? Were many smaller bombs better than just a few large ones? Should delayed fuse bombs be used to destroy emergency personnel and if so, how long should the fuse last?

It also appears that every political leader loved bombing. Churchill bombed the Nazis. The Nazis bombed the British. Stalin then bombed the Nazis from the East. The Japanese bombed the Chinese and the Chinese bombed them back. Then the Japanese bombed America at Pearl Harbor, bringing America officially into World War II.

While I think it's always been implicitly stated by historians, Baker highlights something worth explicitly saying: every side targeted civilians through aerial bombardment. While officially every raid had some justification for reducing industrial output, the bombing raids were also a psychological tool to try and break the enemy's will to fight. Even when spies would report that there was little to no effect on morale, the response from both Hitler and Churchill was simply that more bombers were needed.

From a military perspective, this strategy does make some sense. You risk relatively few men compared to the carnage of trench warfare and can inflict massive damage if you hit your target. Both Churchill and Hitler had served in World War I and the idea of not having a Verdun-style massacre appealed to them both. There was also the fact that the English Channel was once again playing its historical role in making a ground invasion by either side near impossible, therefore creating the massive focus on airpower.

The vignettes on the air war are the one constant throughout the book. If there was an atmosphere created by Baker, it is some combination of air-raid siren and spreadsheet detailing the number, type, and tonnage of bombs dropped that day.

Churchill

Churchill does not come off well, to put it mildly. As head of the Admiralty and later Prime Minister, Winston Churchill's explicitly stated goal was the mass starvation of Europe through international blockade and the installation of abject terror through carpet bombing.

On blockades:

Churchill was the chief obstacle, Hoover wrote later. "He was a militarist of the extreme school who held that the incidental starvation of the women and children was justified if it contributed to the earlier ending of the war by victory."

"When Churchill succeeded Chamberlain as the Prime Minister in May, 1940, he soon stopped all permits of food relief to Poland."

Churchill received countless reports on food scarcity in Europe and did everything in his power to ensure that people starved. This included but was not limited to: blockading food from America and other countries, firebombing farms, releasing invasive insects to eat crops, and spraying an early version of Agent Orange. Churchill was well aware that it would be the

occupied territories that would bear the brunt of the starvation since Nazis would confiscate any food for the war effort. Churchill cynically viewed this as a positive development, believing it would make it more likely that people would rise up against the Nazis. This never happened.

Baker also goes to great lengths to show that Churchill was not a creature of his time or that he was somehow just following the advice of his generals. Once he was elected Prime Minister, Churchill ran the United Kingdom somewhat like a mini-dictatorship, giving no quarter to what he saw as “murderous Huns” being led by Hitler, a “wicked man, the repository and embodiment of many forms of soul-destroying hatreds.” He was a hands-on manager in all aspects of the war, directing everything from troop movements to negotiations with the Americans to the smallest details on weapons manufacturing.

Churchill on the Thickness of Shells:

The thinner-shelled , stronger-blasting German bombs had a charge-to-weight ratio of fifty-fifty. “These are not only more efficient for destroying cities,” Churchill wrote, “they are also cheaper.” Perhaps, the Prime Minister suggested, the charge-to-weight ratio of British bombs ought to be reconsidered - “especially now that the Air Ministry have asked for such a large increase in output.”

Churchill truly was a man of detail. And a political genius. And an alcoholic. And someone who believed that the war must be won at any cost, even if it meant millions would die.

Roosevelt

If Churchill comes off as a warmonger, Roosevelt is portrayed as more of a traditional two-faced politician. When running for reelection in 1940, he repeatedly said that he would not lead America into war.

Roosevelt on Entering the War:

“To Republicans and Democrats, to every man, woman, and child in the nation I say this: Your President and your Secretary of State are following the road to peace. We are arming ourselves not for any foreign war.”

This, of course, was a lie. FDR had begun the mass industrialization and armament of America years before and was actively selling everything from spam to B-24 Liberator bombers to the Allies. Sure, some isolationists and pacifists viewed this unprecedented military buildup as a sign, but who are you going to believe, FDR or your lying eyes?

He had a problem, however. FDR needed what we would now call a “pretext” for war.

On Entering the War:

Henry Stimson was writing in his diary. He, Knox, Stark, Hull, and Marshall had been in the Oval Office with the president, battling around a problem that Roosevelt had brought up. "The Japanese were likely to attack soon, perhaps next Monday," the president said. "The question was how we should maneuver them into the position of firing the first shot without allowing too much danger to ourselves," Stimson wrote. "It was a difficult proposition." It was November 25th, 1941.

FDR was slightly off; Pearl Harbor happened roughly two weeks later on a Sunday. Now I want to be clear since this conspiracy theory territory. There is no evidence that FDR knew the location of the attack on Pearl Harbor or when it would happen. Given the amount of damage to the American Navy, it is hard to believe that he would not have moved at least some ships to safety. FDR in no way, shape, or form allowed the attack on Pearl Harbor to happen.

It is also fairly clear, however, that FDR wanted America in the fight. He knowingly cut off oil supplies to Japan and would then sail tankers just off the Japanese coast. He was smuggling every known armament to Churchill in any way he could. I think it is safe to say given Henry Stimson's first-person testimony and other actions FDR took, he wanted at least one overt attack against American shipping to give him the political cover necessary to declare war.

The fact is, FDR had already tipped his hand by retooling and transforming American industrial capacity, financing, and agricultural power. Even when the Allied countries bled their coffers dry, FDR then pushed extensively for the Lend-Lease Act, essentially allowing unlimited credit to be given for the purchase of American supplies over an undefined time period. The repayment method was left up to Roosevelt by any method he deemed satisfactory. As they say in finance, a nice deal if you can get it.

When the Japanese finally attacked, America was already on a war footing. Perhaps most telling was Churchill's read of his American counterparts on hearing the news.

Churchill on Hearing about Pearl Harbor:

"They have attacked us at Pearl Harbor," said Roosevelt. "We are all in the same boat now..." Both Americans, Churchill observed, took the news "with admirable fortitude. In fact, one might almost have thought they had been delivered from a long pain."

Roosevelt had always known the score when it came to American involvement both pre and post-Pearl Harbor. He coordinated production, shared intelligence, and did everything in his power to both support the Allies and bring the American people around to the idea that once again, America would have to involve itself in a world war.

Pacifists

If the Allied commanders weren't the good guys, then who was? In the afterword of the book, Baker writes:

I dedicate this book to the memory of Clarence Pickett and other American and British pacifists. They've never really gotten their due. They tried to save Jewish refugees, feed Europe, reconcile the United States and Japan, and stop the war from happening. They failed, but they were right.

It's the last sentence which is doing a bit too much work. While they certainly tried to save Jews, did send aid to Europe, and made diplomatic overtures to Japan, I don't know if the world would have been better served by pacifists enforcing US military isolationism. It is always hard to argue a historical counterfactual, but with no United States intervention, Japan would likely have a large East Asian empire based on Japanese ethnic superiority. In Europe, either Nazi Germany survives along with fascist Italy or the Red Army wins and Stalin runs continental Europe. None of these scenarios seems particularly enticing. On the other hand, we do have the benefit of hindsight and now know that Japan, Italy, and Germany are all free democracies. This seems materially better.

The other surprising aspect of the pacifists was just how deeply some held their beliefs. I had always wondered what Gandhi would have done if he were born in Nazi Germany. Well, I now have my answer courtesy of Baker.

Gandhi on Pacifism:

Gandhi answered a letter from Hyim Greenberg, who edited the Jewish Frontier, a liberal Zionist newspaper in New York. Greenberg pointed out that in Germany, A Jewish Gandhi would last about five minutes before he was executed.

"That will not disprove my case," Gandhi replied. "I can conceive the necessity of the immolation of hundreds, if not thousands, to appease the hunger of dictators." The discipline of nonviolence- ahimsa - worked most efficaciously in the face of terrible violence," Gandhi said: "Sufferers need not see the result in their lifetime." It was May 22, 1939.

If we are going to be honest about the moral failings of the Allied commanders, I think it is also fair to point out the failings of the pacifists. When Indians laid down on railroad tracks, the British stopped the train. If Jews had laid down on German tracks, the conductor would have sped up. Given all the horrors that Gandhi knew at the time, it's hard for me to see how nonviolent protest would have stopped Hitler. Extreme pacifism in the face of extreme violence simply does not work. If anything, the Jews of Europe were too pacifist in their reaction to Nazi Germany.

Nazis

The Nazis didn't just loathe the Jews, but looked at European Jewry almost like an engineering problem that needed to be solved. The Nazi high command, led by Hitler, really was megalomaniacal in their focus. Judaism was not a useful political foil to gain power, or a scapegoat, or any other story that has cropped up in recent years by far-right apologists. The Nazis, and the wider German population, truly hated the Jews. Most surprisingly was how organized the hate was - there were rules and procedures that needed to be followed, money that needed to be earmarked, and personnel allocated. You don't just get a Holocaust, it needs to be built.

On Kristallnacht:

Party leaders called their subordinates, and the Gestapo sent out, by Teletype, rules to guide the rioting throughout Germany that was to be the consequence of Ernst vom Rath's assassination. It was to be savage but orderly. The burning of synagogues was permitted "only if there is no danger of fires for the neighborhood." Jewish homes and businesses "may be destroyed but not looted." And foreigners "may not be molested even if they are Jews."

Organized mob may sound like an oxymoron, but Nazi society really was controlled to an extreme extent. The Nazis had a problem, however, in that as they conquered more territory, the number of Jews under Nazi rule actually started to increase.

Thus began a series of experiments.

On Sterilization:

"One practical way of proceeding," wrote Brack, "would be, for instance, to let the persons to be treated approach a counter, where they could be asked to answer some questions and fill in forms, which would take two or three minutes." During that time, their gonads or their ovaries would be subjected to a high-radiation dose emanating from two hidden X-ray sources.

This method was ultimately looked at as too costly and too prone to discovery since the other skin around the genitalia would also become affected by the radiation. Furthermore, reducing the Jewish population over decades through sterilization was simply too slow for Hitler.

On Gassing:

Albert Wimann, a forensic chemist, and Arthur Nebe, the commander of an SS squad, went to an insane asylum and put two pipes through a bricked-in window. When the room was full of patients, they connected an idling car to one of the pipes. The patients didn't die.

Traditional methods of gassing didn't meet the lethality rates that Nazis needed, so more efficient forms would need to be invented. The SS also tried more tried-and-true forms of killing, like mass graves.

On Mass Graves:

Ten people were ordered to get in one of the ditches and lie down. They were told to lie in opposing directions - head, feet, head, feet. And then the policeman threw in grenades, which exploded. Pieces of bodies flew up. Anyone who was still alive after the grenades was shot. The men sprinkled lime and more straw, and the next group was made to lie down on the first layer of dead, blown-up people. More grenades exploded. In Lublin, Poland, that was how 450 Jews died - bombed by grenades at close range. It was October 1941.

Psychological issues started to crop up as soldiers had trouble shooting starving men, women, and children. It was also still not an efficient way to kill the millions Hitler wanted gone. Lastly, it left too much evidence for the Nazi High Command to feel comfortable.

The Nazis continued to refine the killing process until it was used on an industrial scale to achieve Hitler's Final Solution, resulting in the death of over 6 million Jews.

European Jews

Despite what you may have been taught or heard in popular culture, the United States and Great Britain did not fight World War II to save the Jews. No country wanted them. None. Anti-semitism was also so not a uniquely German trait.

Anti-Semitism In Lithuania:

When Grolsh reached the far side of the river, he saw bodies hanging from the trees. A Lithuanian local explained that the people had already "taken care of things": All the Jews in the town had been robbed and hung by fellow Lithuanians. "They had exploited the situation," Grolsh saw. "Hitler is against the Jews anyway. We'll kill them and then we'll take all of their stuff."

It wasn't just Lithuanian peasants who hated the Jews, however. By today's standards, Churchill sounds very much like an anti-semite. I want to give him the benefit of the doubt in that he usually would focus on Communist Jews, rather than the whole religion, but it does appear that he did not have a great love for Judaism. What we do know for a fact is that Churchill turned away millions of European Jews seeking asylum. We also know that of the Jews that made it to Britain, most were eventually rounded up and put into camps because Churchill believed they might be foreign infiltrators.

Churchill on Jewish Foreign Nationals:

Telegrams went out to the chief constables of the English county where German paratroopers might land. It was May 11, 1940. Churchill wanted German and Austrian aliens locked up. Hundreds of people, then thousands, most of them Jewish refugees, were marched by soldiers with fixed bayonets to prison.

There was never any evidence against any of the prisoners, just a general fear because they were German. Of course, Churchill was in charge of an empire where the sun never set, so surely they could find room somewhere?

Churchill on Palestinian Jews:

A group of Jewish refugees from Germany and Austria were put on a boat in the harbor of Haifa, Palestine. With Churchill's approval, British forces were deporting the Jews from Palestine, where they'd arrived illegally after many hardships.

Yes, you read that correctly. Churchill was *deporting* Jews from Palestine. Places as far away as Madagascar and Siberia were looked at as possible homes, but no country was willing to cede any territory or take in millions of refugees. This includes the United States.

Roosevelt on Immigration Quotas:

"They ask that the immigration quota of German Jews to this country be increased from 2,500 to 5,000. This, of course, is almost a negligible number."

Roosevelt's stiff reply - drafted by the State Department - said that there was no immigration quota for "persons in the class described." The State Department had, however, issued 5,117 immigration visas to natives of Germany in 1935: Felix Warburg's request was thus already granted.

Five thousand per year is a shockingly low number given the size of the United States. This policy was not something that was an afterthought, and FDR had numerous opportunities to reverse course, or at the very least take in more Jews.

Roosevelt on Relocation:

A reporter asked if he felt that there was any place in the world that would be able to take a mass emigration of the Jews from Germany.

"I have given a great deal of thought to it," said the president.

"Can you tell us any place particularly desirable?" the reporter asked.

"No," the president answered, "the time is not ripe for that."

Another reporter asked the president if he would recommend a relaxation of the immigration restrictions so that Jewish refugees could come to the United States.

“That is not in contemplation,” said Roosevelt. “We have the quota system.”

Perhaps if the adults can't be saved, at least the US would take in Jewish children, right?

Roosevelt on the Wagner-Rogers Child Refugee Bill:

A member of the House of Representative, Caroline O'Day, tried to reach President Roosevelt to ask him what he thought of the child-refugee bill, which was still alive in committee. Roosevelt's secretary passed on O'Day's message. Roosevelt wrote “File No action FDR.” Without his support, the bill - and the children - had no chance. It was June 2, 1939

It's important to note that FDR was very well aware of the virulent antisemitism of the Nazis and of mass atrocities being committed against them. According to the Holocaust Museum, “Approximately 125,000 Germans, most of them Jewish, immigrated to the United States between 1933 and 1945.” That's a little more than 10,000 per year, a drop in the bucket compared to the total population of European Jews. No matter how you look at it, the United States, the United Kingdom, and the rest of the countries of the world shut their doors to the Jews, knowing full well what would happen to them.

Why You Should Read This Book

It's important to confront uncomfortable truths. History is rarely black and white and when we hide that history, it makes our future decision-making less accurate. It is a good thing to change one's mind, and it's hard to overstate how much this book changed my mind when it comes to Churchill and Roosevelt. Both were given ample, repeated opportunities to save millions of people and chose not to. The amount of human flourishing that was snuffed out is truly unimaginable.

It's also hard to read this book and not think about current events. I'm not going to tell you how to feel or what to believe when it comes to aerial bombing, or immigration, or even war itself. I will tell you that reading this book adds both depth and color to those issues in a way just reading a newspaper or watching television simply can't. For that alone, it is worth the price of admission.

On the Title:

The title comes from Franz Halder, one of Hitler's restive but compliant generals. General Halder told an interrogator that when he was imprisoned in a concentration camp late in the war, he saw flakes of smoke blow into his cell. Human smoke, he called it.

Last Tsar. Nicholas II, His Reign & His Russia by Sergei Oldenburg

Briefly: it is a very comprehensive biography of the last Russian Tsar, Nicholas II. Also, since it is hard, impossible, even, to disentangle his life from the last 20 years of the country he ruled, it is also historical research: hence the title.

Basic information (taken from [here](#)):

The English language edition was published in 1975 by Academic International Press in Gulf Breeze, Florida. Of particular note is the 18-page introduction Searching for the Last Tsar by Associate Professor of History Patrick J. Rollins (now deceased) of Old Dominion University (est. 1930), a public research university in Norfolk, Virginia. As Rollins notes in the study's preface:

“Oldenburg’s [Last Tsar. Nicholas II, His Reign & His Russia] is a major document in modern Russian historiography. The final contribution of a Russian nationalist historian, it provides uniquely sensitive insights into the character, personality, and policies of Russia’s last tsar. It has no rival as a political biography of Nicholas II and is without peer as a comprehensive history of his reign.”

His comprehensive study of Nicholas II is apologetic in nature. Oldenburg substantiates that the revolution interrupted the successful progressive economic development of Russia under Nicholas II: “in the twentieth year of the reign of Emperor Nicholas II, Russia had reached an unprecedented level of economic prosperity”.

Oldenburg was able to undertake such a study of Russia’s last tsar, having had access to a unique collection of documents. These included copies of authentic historical acts of the Russian Empire held in the Russian Embassy in Paris on Rue Grenelle. Long before the First World War, duplicates of the originals had been made as a precautionary measure, and sent to the Russian Embassy in Paris for storage. In October 1917, the Provisional Government appointed Vasily Alekseyevich Maklakov (1869-1957), to replace Alexander Izvolsky as Russia’s Ambassador to France.

When he arrived in Paris, Maklakov learned about the takeover by the Bolsheviks. Regardless, he continued to occupy the splendid mansion of the Russian embassy for seven years, until France found it necessary to recognize the Bolshevik government. Fearing that the Embassy’s archival documents would fall into the hands of the Bolsheviks, Maklakov packed them up, including Oldenburg’s manuscript, the Okhrana archives, among other items and arranged

for their transfer to the Stanford University.

Oldenburg's fundamental historical research on the life and reign of Emperor Nicholas II, is sadly overlooked or simply ignored by Western historians.

First, yes, Oldenburg's research is really thorough and comprehensive. But I would say all the unique official documents and data are not the only things which make the book great: whatever the issue discussed is, he extensively quotes the positions of the whole Russian society about it—from every part of the political spectrum and every group of interest. In a sense, you regularly get a detailed cross-section of the whole vast country. This book is a masterpiece and should be a required read for anyone who wants to understand the period.

Second, indeed, it's interesting to note that despite the quality of research in the book, it isn't really any popular in the West. The most popular edition on Amazon occupies honorable ~5 000 000-th place in the bestsellers ranking. As noted in [this master's thesis](#):

The only work which has discussed with any depth the reign of Nicholas is Sergei Oldenburg's work, *The Last Tsar*. Originally published in two volumes in Russian in 1939, this work has received little attention by historians.

...

The work is available in a four volume English translation, but is either ignored or totally unknown to historians. In the introduction to the first volume of Oldenburg's work, editor Patrick J. Rollins observed: "The real emperor, according to Oldenburg, was a strong-willed, independent minded monarch who personally directed Russia's foreign and domestic policies and who took counsel only with himself." But for some obscure reason, Oldenburg's work is used primarily—best exemplified by W. Bruce Lincoln—to recount Nicholas' fears at his succession to the throne, his 'senseless dreams' speech to the zemstvos which historians have emphasized as best portraying Nicholas' personality, and his Abdication Manifesto which is available in almost all collections of documents of the era.

It's way too easy to explain it away as "snobby liberal historians hate Nicholas and don't want to take a look at anything which contradicts their narrative". However, I'm pretty sure the circumstances of the book writing play at least as huge a role here. Are there many relatively popular Russian historians in the Anglosphere? The only one who I could think about might be Lev Gumilyov, and that must be in large part due to his and his father's tragic life (and death) in the USSR; also, curiously enough, as a response to

"last tsar" query Amazon manages to show me something by Radzinsky. But "Russian émigré-historian" is someone who belonged to a small and scattered community which wasn't backed up by the resources of any nation-state. However good your books are, any popularity contest is going to be an uphill battle in such a condition; this battle was probably won only once by Nabokov.

Of course, there are many things in the book. The first years of the Emperor's reign; the Russo-Japanese War; the Revolution of 1905, which didn't dethrone the Czar, but made him establish the Duma (parliament, pronounced as "doom-ah"); the agrarian reforms; the First World War, but...

But the real meat for me—of course—would be the circumstances of the Revolution of 1917 and the people who created it.

Imagine you're watching a horror movie. The producer was really lazy, and the movie feels like nothing but just a bunch of hastily lumped together cliches. A bunch of drunk teenagers wanders outside the city, goes *The Ways You Are Not Supposed To Go*, enters *The Abandoned Building You Are Not Supposed To Enter*, reads some *Old Books In Forbidden Languages You Are Not Supposed To Read* (all this while a couple of sober and level-headed characters scream as much as they can "JUST WHAT ARE YOU DOING, LET'S GO HOME WHILE WE CAN"), and then they summon some ancient entity which tortures, devours and inflicts other fates-worse-than-death to these teenagers and everyone and everything they ever held dear.

Those are basically the parts of the book related to the build-up to Revolution and the revolutionaries. Well, not quite—it stops right after the Revolution and the Czar's abdication (no surprise, given the title). But it certainly gives the reader the taste of horrors to come, and it's not like I haven't known what happened next.

The description of revolutionaries reminded me a lot about that de Maistre's idea about how the French Revolution was a culmination of Enlightenment—and, hence, a kind of a providential divine punishment for all the cool and fashionable people of France who cheered for Enlightenment and republicanism in 18th century and who got murdered by it. This in a sense certainly applies to the Russian Revolution too. The people who wanted to remove the Czar were playing with fire, but they thought this fire was a small cool firecracker; it turned out to be a (no pun intended) Czar Bomb, obliterating all their plans and strategies in a heartbeat.

A couple of good examples: 1) general Ruzsky who actually convinces the Czar to abdicate (while having the—I don't know—, gall or insanity to claim that it's *actually* for the Czar's and his dynasty's own good and safety, of course) gets murdered by the communists the next year; 2) the head of the "progressive Bloc" in Duma Guchkov who spent most of the 1916 working hard on creating the revolutionary conditions, *after just the first ~week of the Revolution* still cheers the dethroning of uncool monarch, and yet when he sometimes tries to understand who is actually in charge, he is like "wait wait wait, hold on, no, no, not like this, Oh God, NO!!!"; needless to say, it only gets worse from there.

These examples are not an exception, they are more like an iron rule: not a single "moderate", "constitutional", "liberal" revolutionary in the Army or in the Duma gets what he wants (the most lucky ones were those who at least were able to escape to the West); to remember that famous Mencken quote: they certainly got what they deserved, and they got it good and hard.

But why did these people cheer for the Revolution? Why did they all want—the things they wanted?

Contrary to all the narratives about how it's only dumb people who believe dumb things—probably because they are not educated enough, people of any intellect can be captured by all kinds of fads (or dare I say, *toxoplasmas*?). Moreover, since these fads often are functions of not only IQ, but social standing, religious beliefs, and political persuasion... Yeah.

The hate intelligentsia felt towards the monarchy is just that: a fad. "the Czar, is, like, boring and uncool, but democracy and socialism are pretty lit". And this fad certainly doesn't appear to me as arising from pure search for truth, free from any piece of self-interest. Anyway.

During the Revolution of 1905 and subsequent mayhem they tried to do their best to squeeze any concessions from the uncool Czar (ideally, of course, to make him resign). The Czar can't really find enough societal support for him to continue being the sole Ruler, so he has to concede a bit and establish the Constitution and the Duma. Duma doesn't have much power: it's just a legislative body, and Nicholas reserves for himself the powers to suspend it, enact laws without its support, and declare the re-elections. Of course, these are the powers of the exceptional circumstances, not to be used on a day-to-day basis. He uses them several times (mostly in the appropriate moments; once

—maybe not exactly, probably asking him to do it that time was the only mistake of his PM Stolypin, otherwise an amazing statesman). However, he doesn't do much to Duma during the wartime, which certainly feels like a mistake (cf. e.g. Habsburgs in Austria, who just cancelled all the elections and the Reichstag in 1914).

This system wobbles its way through 1905-1914, sometimes producing completely disastrous results, sometimes—somewhat decent ones, when the Duma doesn't have too many batshit insane revolutionaries and its members are in the good mood. If this is any improvement over the good old autocracy, I'm a ballet dancer.

The interesting thing about this structure is that it looks like a disaster pretty much regardless of your political positions. Democracy—politics—populism tends to select for the worst kind of statesmen; this was as true in 1916 as in 2016. The right wing of parliament was foolish, chauvinist and vulgar; the "supporters of constitution" wing, the ones who apparently wanted to subject the government to the rule of law, didn't really care about *the actual laws*; and the left wing instead of helping workers or peasants was just mostly excited about "property redistribution".

Then the war happens. I'm... really not sure of what to think about its beginning. I remember reading an opinion that Serbian state apparatus was complicit in the Sarajevo murder, and Habsburgs had the perfect right to retaliate. Anyway, Oldenburg doesn't give too much attention to the Sarajevo incident (and he seems to think Russia was right in helping Serbia—although Russia formally was not even a Serbia ally!).

However, it doesn't look like the Czar had much choice there. Curious enough, no matter how strong Russian intelligentsia hated their Czar, they hated the Central powers (German Empire + Habsburgs' monarchy) much more. These nations were even less cool, they were even—for the intelligentsia it was the greatest insult— *reactionary*.

And Russian intelligentsia was strongly Anglo-philic, and the United Kingdom kind of was going to take (and took it alright) the anti-German side.

So, yeah, for the first months of war everyone is crazy, liberals forget about their hate for the Czar and unite with him to smash the hated Habsburgs and Hohenzollerns (but also to gain more political power through controlling the war industries; which somewhat happens). Right-wing Slavophiles jump the bandwagon too to help Slavic nation of Serbia. There are some reasonable people who predict how incredibly disastrous this whole endeavor will end up; they certainly appear to be in the minority.

Then the war doesn't end overnight, and everyone becomes kinda jaded. The liberals realize that they won't be able to defeat Germany any time soon, so they forget about it

and start to work around the clock to bring the Czar down though any kind of slander and lies they can, using the Duma as a bully pulpit and utilizing their connections from the war industries to go around the censorship (which is—like all the imperial apparatus—meek and vegetarian, and much more liberal than in e.g. France; the communist nickname of Nicholas "Bloodthirsty" looks like a fantastic case of projection). A lot of "right-wingers" and "monarchists" start to defect to the liberals' side and spout their lunacy (ffs, what was wrong with someone like Purishkevich to actually pull something like that?).

Then in early 1917 the Duma's term is about to end, and the Czar is thinking about not convening it again, but before this happens, Petrograd — i.e. Saint Petersburg, which at the time was Russian capital — becomes engulfed in bread panic which then becomes bread riot (this is more or less the first time there were any major issues with food there: that the food supply in the Russian Empire was pretty much way ahead of every other major warring country is, like many Russian Empire-related things, an acknowledged but overlooked fact). The riot is suppressed, but then immediately — after months of revolutionary propaganda—an army regiment mutinies too. This mutiny is suppressed as well, but several others flare up, and soon Petrograd is controlled — if this word is applicable here — by the soldiers. The Czar tries to move the loyal regiments to the city and get there himself, but he is stopped by either dumb or Revolution-sympathizing general (see above, also, while I can believe he was thinking that the Duma was controlling something (in reality the true power in the city at the moment has already belonged to organizations like the "Councils of Revolutionary Soldiers"), I'm not sure if he was dumb enough to think that the Duma controlling city was actually something good, given what the Duma thought about the Czar; anyway...). This is the early March of 1917, and that's how the Empire ended: not with a bang but a whimper.

Right now almost every country in the First World is a liberal democracy (even if it has some decorative royal figure). Some centuries ago almost all of them were absolute monarchies; at some point most of them were both monarchies and democracies at the same time; there were two centers of power: the monarch, and the electable body. These transitions were often slow and peaceful (like in e.g. England after 1688), and the reasons behind them may be not exactly clear. Curtis Yarvin once wrote that if you had claimed e.g. in England after the Glorious Revolution that this arrangement was unstable, you would have been instantly labeled as an ultra-right "Tory" absolutist-monarchist — but the history would have slowly proved you right. The history of representative democracy in Russian Empire demonstrates very well why this happens: you can't really have more than one center of power. They will try to compete with and subjugate each other until only one such a center will remain. In England this center

turned out to be the Parliament which defeated the Crown. In Russia the Duma was pushing, pushing and pushing, never capitulating completely and always preparing to pump the pressure up again whenever possible. Probably the most serious issue with the Czar was that he didn't have enough guts to play as hard a ball as the situation required. Not so seldom he was trying to make a compromise with the Duma by appointing more liberal ministers. Needless to say, unless these ministers were following the Duma's positions rather than Czar's, they were obstructed and boycotted: the Duma never gave any quarter to the Czar. While, in my opinion, the very least every single member of "progressive bloc" deserved was lustration, probably not so radical methods could have helped too — at least, for a while. In short, so much for the "checks and balances". Thank you Montesquieu, very cool!

It's interesting how you can notice this pattern even in temporary political struggles and positions of people who gained power: in 1917 no one needed moderates except as temporary useful idiots. Two stable positions were "I trust the full power in Czar", and "I trust the full power in People". "I trust a bit in Czar, but he should give some way to People" isn't really a way to win friends and influence people: it's a way to gain some temporary alliance which will end when the radicals will make your anti-Czar arguments more refined and more pure (and probably murder you for being not revolutionary enough).

To sum, great book. It sort of confirms my priors about the Empire, and it somewhat changes them when it comes to Duma specifically and parliamentarism in general: I didn't hold Russian politicians of the era in high esteem, but I didn't expect them to be that bad in the art of, well, managing the state (although they were certainly pretty good in intrigues and manipulating mass opinion!). Again, highly recommend it to anyone; regardless of your political affiliations, if you are even slightly interested in early 20-th century history, reading it most likely will be pretty rewarding for you.

Some random thoughts and quotes:

- It was quite interesting to brush up on the actions of far-left revolutionaries in rioting Petrograd in March. Lenin's October formula "To capture post offices, telegraph offices, and train stations" has been quite famous in Russia, but it was curious to see that socialists even in March had pretty strict contingency plans on the subject of "What to do when the SHTF in Petrograd" (to wit: free their leaders from prisons, telegraph the situation to other cities, ask workers at every factory

to appoint a representative to the Workers' Council, etc., etc., etc., evil those people might be, but their coordination and strategic thinking are certainly quite impressive and instructive).

- Oldenburg gives an interesting quote about the events by Winston Churchill, from "The world crisis. 1916-1918":

Surely to no nation has Fate been more malignant than to Russia. Her ship went down in sight of port. She had actually weathered the storm when all was cast away. Every sacrifice had been made; the toil was achieved. Despair and Treachery usurped command at the very moment when the task was done.

The long retreats were ended; the munition famine was broken; arms were pouring in; stronger, larger, better equipped armies guarded the immense front; the depots overflowed with sturdy men. Alexeieff directed the Army and Koltchak the Fleet. Moreover, no difficult action was now required: to remain in presence: to lean with heavy weight upon the far-stretched Teutonic line: to hold without exceptional activity the weakened hostile forces on her front: in a word, to endure—that was all that stood between Russia and the fruits of general victory. Says Ludendorff, surveying the scene at the close of 1916:

'Russia, in particular, produced very strong new formations, divisions were reduced to twelve battalions, the batteries to six guns; new divisions were formed out of the surplus fourth battalions and the seventh and eighth guns of each battery. This reorganization made a great increase of strength.'

It meant in fact that the Russian Empire marshalled for the campaign of 1917 a far larger and better equipped army than that with which she had started the war. In March the Czar was on his throne; the Russian Empire and people stood, the front was safe, and victory certain.

It is the shallow fashion of these times to dismiss the Czarist regime as a purblind, corrupt, incompetent tyranny. But a survey of its thirty months' war with Germany and Austria should correct these loose impressions and expose the dominant facts. We may measure the strength of the Russian Empire by the battering it had endured, by the disasters it had survived, by the inexhaustible forces it had developed, and by the recovery it had made. In the Governments of States, when great events are afoot, the

leader of the nation, whoever he be, is held accountable for failure and vindicated by success. No matter who wrought the toil, who planned the struggle, to the supreme responsible authority belongs the blame or credit for the result.

Why should this stern test be denied to Nicholas II? He had made many mistakes, what ruler had not? He was neither a great captain nor a great prince. He was only a true, simple man of average ability, of merciful disposition, upheld in all his daily life by his faith in God. But the brunt of supreme decisions centred upon him. At the summit where all problems are reduced to Yea or Nay, where events transcend the faculties of men and where all is inscrutable, he had to give the answers. His was the function of the compass-needle. War or no war? Advance or retreat Right or left? Democratize or hold firm? Quit or persevere? These were the battlefields of Nicholas II. Why should he reap no honour from them? The devoted onset of the Russian armies which saved Paris in 1914; the mastered agony of the munitionless retreat; the slowly regathered forces; the victories of Brusiloff; the Russian entry upon the campaign of 1917, unconquered, stronger than ever; has he no share in these? In spite of errors vast and terrible, the regime he personified, over which he presided, to which his personal character gave the vital spark, had at this moment won the war for Russia.

He is about to be struck down. A dark hand, gloved at first in folly, now intervenes. Exit Czar. Deliver him and all he loved to wounds and death. Belittle his efforts, asperse his conduct, insult his memory; but pause then to tell us who else was found capable. Who-or what could guide the Russian State? Men gifted and daring; men ambitious and fierce; spirits audacious and commanding —of these there was no lack. But none could answer the few plain questions on which the life and fame of Russia turned. With victory in her grasp she fell upon the earth, devoured alive, like Herod of old, by worms. But not in vain her valiant deeds. The giant mortally stricken had just time, with dying strength, to pass the torch eastward across the ocean to a new Titan long sunk in doubt who now arose and began ponderously to arm. The Russian Empire fell on March 16; on April 6 the United States entered the war.

- This moment was not mentioned in the book, but one of the important aspects of Communist propaganda (and "post"-communist, democratic actually too: they are quite synoptic on the topic of Russian monarchy) was "Czar was so disastrous a Commander, that Russian army lost lots of area in 1915". The Army lost a lot of

area indeed: it's not for nothing they called the events of 1915 "The Great Retreat". And yet, if you look at the actual map, you'll see that this land lost (and actually partially recovered in 1916) is just peanuts compared to the vast, enormous extents the Communist army gave away in 1941 (and in much more disorganized manner, with much more severe personnel losses). Talk about throwing stones from the glass houses! (also, to be precise, the Czar appointed himself the commander-in-chief *after* The Great Retreat)

- There's also an interesting piece on "what the monarchy did wrong" from some monarchical address used as the afterword in the book. I read the book in original, so here is my translation of the excerpt:

One of the most important reasons why the power consciously diverged from society was the distinctive policy of using the lower class as a support base over the heads of the middle class. This idea—of contraposition of "good people" and "evil intelligentsia"—a kind of a "right-wing populism"—was familiar to many Rulers, starting with the emperor Nicholas I. And this is the antithesis to the idea The Great Catherine built her power on. This is the policy of Ivan the Terrible against the boyars, or of Gustav III against the Swedish nobles, etc., this piece is not the place to evaluate this policy; and yet both late Emperor and Empress had this thought sometimes; it was somewhat close to Witte too, but to a much lesser degree—to Stolypin.

One of the dangers of this path is that it can easily lead you down the road of losing the understanding of what's real in politics and what's not. Some fictions and conventions start to appear. The living bond between the power and people is replaced by the bond theoretical. The belief that people are kind of real, good, devoted to their monarch, was alive in the Ruler until his last days. He could have been right—but the Czar's voice couldn't get to people through the inimical atmosphere. And so the Ruler had the distorted image of society and people, and people and society had the distorted image of the Czar.

Eschewing any true information about the Czar and his family, Russian intelligentsia was extremely biased and eager to listen to what underground revolutionary lampoons wrote about him—usually it was just extremely lunatic fantasy, to whispers of courtroom gossips, to insinuations of disfavored dignitaries. The belief that the Ruler was ignorant, narrow-minded—some even dared to say "retarded"—that he was a man lacking in the will, and yet, evil and insidious—was quite

common among intelligentsia. Even his military rank—which He was stuck in, because His Father died when the Ruler was 26 years old—was used to rebuke him, when they were speaking of "small colonel", of his "rank"—for some reason—"of army colonel", etc. That's how there appeared the atmosphere between the monarch and people which didn't let any good through and let only evil pass.

(I would say I mostly agree with this, though it's really not obvious to me how to break this vicious cycle of "intelligentsia despises the government => they don't want to collaborate with the government => the government lacks smart people => intelligentsia despises the government even more" without resorting to direct concessions to intelligentsia which, of course, only make them realize that the power is weak and make them want to ask for more, until the sovereign is dethroned).

- The book also demonstrates very explicitly and utilizing many sources just how much most of the talking points about the Revolution we still repeat (e.g. the Rasputin story) were pretty much political propaganda created hastily just for the purposes of taking power. Not that I didn't know it, and yet: kind of sad.
- Another curious thing: lots of parliamentary quarrels, animosity and disagreements are just incredibly... petty? Going back to Guchkov: he was a leader of Octobrist Party, basically moderate monarchists/right-wingers; the Czar thinks of him well enough to give him a certain prominent position; after that they two have a private conversation, which Guchkov then partly leaks to the media; the Czar is really pissed off, and Guchkov goes full "well, then screw you too" and later leads the effort to force the parliamentary sovereignty on the Czar. This is not how the government of any country should look like, this is some kindergarten-level crap.

Loserthink by Scott Adams

If a comic-strip writer wrote a book with a certain tagline, you shouldn't be too surprised if there is a practical joke in there. If the writer is famous for one of the most successful comic strips of all time, which might co-incidentally rhyme with Shill-hurt, then I'd say the odds of that go up.

The tagline of Loserthink by Scott Adams is "How Untrained Brains are Ruining America" and is set in a time of increasing political polarization. But the book has very little to do with the details of the American political landscape. I don't think this is Adams way of a practical joke (although it might be), but far more likely that it is the book tagline equivalent of clickbait.

Loserthink is a great book and is somewhere on the spectrum between self-help and social commentary. If I had to put a number on it, I would say that it's close to 70:30, in favour of the self-help component. But this is based on memory, and while I have a fairly decent memory, all human memory can be surprisingly flawed. In any case, the personal components of Loserthink are far more interesting than the political and current events discussed in the book. Further those events are now outdated anyways and are of no use to anybody. Accordingly, in this book review I have only covered the points which are timeless and will likely continue to remain relevant moving forward. All facts, politics, trends (including the ultra-positive golden age trend) have been perfectly ignored because they are of absolutely no use to anyone anymore.

One technique that Scott Adams often talks about in relation to persuasion is framing. It's a powerful one. The exact same set of facts can be look completely different depending on the frame you place on it. A few examples of this:

1. Peter Thiel donates a few million dollars to life-extension.
 - a. Effective Altruist: Peter Thiel is extremely altruistic and is willing to spend his money in risky ways that can benefit humanity in the long run.
 - b. Skeptic frame: A business savvy Peter Thiel got lucky with the internet, his networking skills and his financial wizardry to work on the right products, but is now squandering his wealth on projects that will never yield returns and this can only be done by someone who does not understand science and in particular biology.
 - c. Venture Capitalist frame: Thiel is investing in several high-risk high return ventures and he knows that if even one of them such as life-expectancy pays off, he will be many times richer in the future than he is now and will be a super successful venture-capitalist, perhaps the greatest of all time.
2. Elon Musk starts a rocket company to go to Mars, a car company and a tunnel company.
 - a. Engineer frame: Elon Musk is an engineering genius. The projects that he takes on are so grand and complex that his motivation must be to challenge his engineering prowess. This is inborn talent, the likes of which normal people cannot compete with, even though some people may develop expertise in some specific areas after decades of study and work.

- b. Chosen One frame: There exists a prophecy that one man will save humanity by building rockets. He will combat global warming by making better cars and tunnels along which the cars can move.
 - c. Workaholic frame: Elon knows that he needs to put in long hours at the job or else he will go crazy. In order to stop him from having a soul crushing existential crisis, Elon needs to take on projects that are so grand just to keep his mind occupied and get his work fix.
3. Eliezer Yudkowsky writes fiction and a comprehensive blog on rationality
- a. Genius frame: Yudkowsky is a genius and realized that some of his musings on human thought process is not just useful for his AI-themed work, but can also benefit others. A casual blog post for him might be life-changing insight for one of his readers and it often is.
 - b. Big mouth frame: Yudkowsky is very verbose and despite having a complicated subject that he works on, cannot resist the urge to stop talking and/or pen down his thoughts on a blog. He knows he's smarter than 99.99999% of humanity but keeps writing his most obvious thoughts just for the love and attention he gets from us lesser beings.
 - c. Process frame: Even geniuses need process and having too many thoughts in your head at the same time can be maddening, so writing them down in an organized series of posts can help you get some much needed clarity of thought.

The point here is not to suggest that any one frame is right and another frame is wrong. That would be missing the forest for the trees. It's about realizing that there are multiple ways to think about a single issue/process/phenomenon.

Charlie Munger, one of the most successful investors on the planet, has a phrase for this. He calls them "mental models". And he attributes much of his success to be able to look at the same set of facts through different mental models.

If I had to describe Loserthink in one sentence, it would be "A book about using different frames or mental models to make decisions and take action". That would make a much better tagline than some quip about America.

The important lesson from Loserthink is that recognizing which frame/mental model should be adopted for which situation is a skill in itself. And it's a meta-skill which only gets developed once you have the underlying basic skill-set.

Some of this may sound intuitive if you are an action-oriented person that has dabbled in multiple disciplines. Otherwise, it might also sound slightly hand-wavy akin to an eastern mystic saying something along the lines of "All truths are equally valid." Indeed, the big problem with much of philosophy books and business-advice books is not the advice that you find in those books, but knowing when to apply such advice.

One striking example of this problem from philosophy is the advice “If you meet the Buddha on the road, kill him”. As pointed out by Naval Ravikant, what this is intended to mean is that if you are in search of enlightenment, there are many states of mind that look like true enlightenment, but are in fact not.

Loserthink has some practical ideas around this problem. Largely, the book is organized into many chapters titled “How to think like a _____”. The blank space is filled by (1) Psychologist; (2) Artist; (3) Historian; (4) Engineer; (5) Leader; (6) Scientist ; (7) Entrepreneur; (8) Economist.

At this point, if you’re thinking something along the lines of “Yeah, we get it. Separate magisterium and all that”, then I urge you to show a little more patience because Adams doesn’t keep all these chapters completely unrelated to one another.

He simply uses each of these professions as proxies for different ways of thinking, illuminating that each of these professions are aware of some secret about how the world works, which are not obvious to persons outside of that profession.

I’m not the expert, but I imagine that readers of this blog are largely familiar with Bayes theorem. It is hard to argue a better framework for truth-seeking. The universe is truly Bayesian.

However, success in life and truth-seeking prowess are poorly correlated (with the exception of a few specific domains). Loserthink is a book about how to take action and how some frames are more suitable than others depending upon the situation.

Think like a Psychologist

LessWrongers often talk about the illusion of transparency. A much bigger problem is the reverse, where people assume they can read others minds. Adams then goes on to attribute much of the polarization in the world to specifically this phenomenon. While Adams concedes that people don’t *literally* believe that they are reading another person’s thoughts, he does believe that everyone falls for this illusion. People believe based on out of context speech that even a simpleton could see [insert apparently obvious] intent.

Similarly, Adams goes on to identify the issue of psychological projection as a real problem but one that untrained individuals won’t be able to identify.

Adams talks about embarrassment and lays down strategies to deal with embarrassment, specifically learning to actively notice how fine you are after an embarrassing incident. Conversely, Adams recommends noting how little other people’s embarrassment mean to you. Usually this amounts to nothing. Adams then proceeds to mention that using the above techniques he went from being embarrassed about everything to having no shame whatsoever.

Another important lesson from this chapter is about noticing wrongness. This one is super important from a happiness point of view. Adams takes an example of how he notices that one end of his tablecloth is wrinkled and he is not able to move his attention away from that until it is resolved. He then goes on to marvel at how despite the fact he is wealthy, has fame, has a beautiful significant other, and generally has other things going for him in life, he is unable to shift his attention away from this one thing. This is Adams' way of driving the point home that we need to count our blessings in life. We really do have limited shelf space in our minds. It is important to choose wisely what we do with it.

Most importantly from this chapter, Adams talks about ego, but I consider that important enough to warrant its own section in this book review.

Think like an Artist

According to Adams' one of the defining features of an artist is their imagination. In general terms, I agree with this. The entire crux of this chapter is essentially this. Adams offers the parting words at the end of this chapter: "If you can't imagine any alternate explanation for a particular set of events, it might be because you are bad at imagining things"

It's important to note that Adams has consistently maintained in public that the individuals least capable of critical reasoning are often those trained only in the arts. This is not inconsistent with his world-view that cross-disciplinarily trained minds are superior.

Think like a Historian

Adams makes two important points about history. The first is that most history is probably distorted due to the incentives of the people writing it. This includes both that history is written by the winners, but also that history is taught in different ways in different countries.

The second point is more profound. History can often have a disproportionately strong hold on us. I have personally struggled (and continue to struggle) with this. At such times, it is important to remember that everything that we call history is in fact completely imaginary. It exists only in our minds.

This is best summed up with the following lines: "*Focusing on the past when the present offers sufficient paths to success is Loserthink. It is better to focus on your own systems for success and when you succeed watch how winning fixes most problems.*"

Adams also points out that the phrase is “history repeats itself” functions more like a [stop-sign](#) than as an actual predictor of anything useful.

Engineer

It should come as no surprise that the creator of Dilbert has something positive to say about engineers. Alas, this chapter had less to do with engineering per se as a discipline. As is the case, throughout the book, Adams uses engineering in the headline, to make the point about situations and their solutions.

The two key take-aways from this chapter: (1) the best solution to a problem does not consider whose fault the problem was; (2) most situations are far more complex and trying to point to a single variable for a solution is probably a bad idea that won't work.

Leader

In stark contrast to thinking like a scientist, Adams' notes that a strategy adopted by effective leaders is often to set targets that seem overly ambitious. Adams' does not cite this in the book, but Elon Musk has been known to do this repeatedly. The argument is that the job of a leader is to inspire the troops and motivate them to work hard towards a specific goal.

For such a goal, the truth is often not the most useful weapon. Hyperbole is much more effective in such cases (Sorry, Bryan Caplan!).

Scientist

This is probably the most basic chapter of the book and if you're a regular reader of LessWrong, SlateStarCodex or any similar blogs, then the contents of this chapter are the easiest and most intuitive. You could consider skipping it, but if you're a reader of LessWrong, SlateStarCodex or any similar blogs, you're probably the kind of person who reads every single word without skipping any.

Adams opens the chapter explaining how frequently coincidences occur and why they should be treated with skepticism. Adams then proceeds to give anecdotal evidence the same treatment.

Adams recommends (and I fully endorse this recommendation) of regularly engaging in an exercise of asking oneself what things would look like if the opposite of what you

believed was true. Put in far simpler terms, but somewhat less eloquently than Eliezer does when he is talking about Bayesian evidence.

Entrepreneur

Adams opens this chapter with an amusing anecdote on how to overcome couch lock. If you don't know what couch lock is, it's the feeling you get when you can't get off the couch, probably because you've smoked too much marijuana. If you've never smoked too much marijuana, then you can consider smoking a lot of it just to understand what the hell I mean by couch lock. Alternately you can imagine yourself wanting to do something but completely lacking the motivation to do so. In other words it's a form of akrasia.

Adams' solution (which I recommend) is to start small. Wiggle your finger and eventually get to moving your entire body off the couch. It's fairly standard self-help advice. But it's a massively useful life hack.

Whenever confronted with a large project, stop thinking about the project as a whole as that will paralyze you into inactivity. Think about the micro-steps you can take and focus only on those micro-steps.

Adams has long been an advocate for talent-stacking or picking up multiple skills. In this chapter, Adams elaborates on the process. Adams claims that two of his favourite sentences [to hear in succession] are: (1) I don't know how to do that; (2) but I can figure it out.

Adams also stresses an important principle that forms part of a lot of startup thinking today, namely the importance of being able to test assumptions in a way that no one gets hurt.

Economist

A personal gripe I've recently developed is that whenever I deal with something to do with economics, whether it's a book, or a friend, I get drowned in models. No one ever seems to talk about the most important part of economics – incentives.

As a startup founder, I'm slowly coming to the realization that incentive structures are the most important and yet underrated tools we have to explain how the world works and modify it to suit our needs.

And yet, no one seems to be talking about it enough.

Another part of economics that needs to be spoken about is the issue of scarcity. But just like he did with all other professions, Adams takes out the best part of economist thinking and uses them to illustrate decision making principles.

Adams opens up the chapter by explain when it makes sense to speed and when it doesn't by using reasoning on how the police department are likely to make their staffing decisions.

However, the important point on this chapter is that the people who understand economics and incentives can more easily spot hoaxes, because money can drive human behaviour in predictable ways. As a corollary it pays not to trust experts who have financial incentives that are not aligned with yours.

Adams also makes a point about making comparisons. Most people who aren't trained in economics (or in rationality) do not have the ability to accurately compare things but are left with the illusion that they can with perfect confidence. Comparing things to alternatives is an important skill to develop and this is an important benefit that the LessWrong and it's allied communities receive while practicing Bayesianism.

Comedian

To be honest, Adams doesn't have a separate chapter on this. That's a little surprising since he's a professional funny man.

But a quote from Jerry Seinfeld, did come to mind recently for me. Seinfeld is walking with a fellow comedian (this was a TV show called Comedians in Cars with Coffee) and they are talking about how a comedian's thought process is different from everyone else's.

"You know how when you give your laundry, you have so many options? But when you collect your laundry, you only have one option". Now this isn't a joke in itself, points our Seinfeld. But it is a great example of how comedians think.

I think it's important to learn how comedians think. Firstly, because we could all do with more laughter. But the truth is that comedians have great bullshit detectors. If an economist can detect a scam because he has a deep understanding of incentives, a comedian can detect an inefficiency because he has a deep understanding of wrongness.

There are some other ideas that Adams has woven into the book as well – new, novel ideas entirely of Adams creation which are, in my opinion, the most powerful portions of the book.

Using the Ego as a Tool

In my opinion, this is the most important idea by far. So much so that I pondered writing an entire book expounding on the basic idea myself. I haven't done so, partly due to akrasia, but also partly because I would feel a lot better about writing such a book after my personal accomplishments are sufficient to inspire a non-trivial number of other human beings. Ironically, mastering your ego is enough to circumvent both of those hurdles.

The conventional wisdom on your ego is that one must always look to suppress it. It is also looked at as largely a social element and much less of a personal one. Yudkowsky has rejected this paradigm, [drawn the distinction between fake social modesty and true humility](#).

Adams offers practical advice on exercises that you can adopt to increase or decrease your ego as the case may require. In certain situations, ramping up your ego can give you much needed courage. A good example of this would be public speaking.

Yet all improvement in any domain comes from iterations which necessarily mean admitting your mistakes and making changes. That can only happen with humility as both Adams and Yudkowsky have illustrated.

Upon thinking about the ego as a tool that can be raised up or decreased, I came to the realization that this goes far beyond just what Adams describes in Loserthink. All emotions can be dealt with in one way or another by controlling your ego.

Anger, is nothing but rapidly increasing ego and making a conscious effort to reduce your anger goes a long way in calming down.

Conversely, depression feels like it can be overcome by simply raising your ego. Having a big ego feels good and that's why people do it (even when they are boasting of their modesty).

Discipline, is nothing but ensuring that you don't get over-confident to skip even a single day's work. That's controlling your ego too.

Any entrepreneurial activity such as trying to build a company and a lot of what some of the smartest people in the world have accomplished are largely the result of high intelligence coupled with hard work and high productivity. But none of them would have tried out such ambitious projects without having an ego. Having a large ego allows world-class athletes to put in the extra hours of training, and startups on the verge of bankruptcy to continue with their mission.

What about rationality itself? The LessWrong community and the figures commonly associated with it are all paragons of the virtue of calibrating your confidence level to ensure that you are neither over-confident nor under-confident. Perhaps no one is a master of this more than Bryan Caplan. While the LessWrong community preaches

tying this confidence to what is grounded in verifiable fact, Adams proposes that it can be used at the behest of its owner.

Following this train of thought, it led me to conjecture that perhaps this is the true function of religion. Humans have a need to maintain a constant self-image per Robert Cialdini. This makes it hard to modulate your ego, without deliberate practice.

But the presence of an all-powerful deity who has your back at every turn gives you all the courage in the world when you need it. And at the same time, it gives you all the humility that you need even when things are going perfectly according to plan. While I do not personally know a single religious person who thinks this way, I have a strong suspicion that this is in fact the case and that most religious people have not decodified it to absorb this information.

Conclusion

Adams at one point notes that “Stay in your lane” maybe the advice that he gives his worst enemy. Most progress in life comes from venturing out and performing activities that involve things outside of your comfort zone. This dovetails well with another piece of advice that Adams has now become well-known for: “The Talent Stack”.

In Loserthink, Adams lists out that he has developed skills in economics, business, sales, management, psychology, hypnosis, programming, commercial lending, project management, public speaking, design, art, writing, television production, negotiating, budgeting, persuasion, marketing, social media, video editing, engineering, drumming, cartooning, political punditry.

An important point that is worth mentioning: Adams throughout the book refers to various mental traps as being stuck in a mental prison. It is a fitting metaphor. One that even Kanye West approves of.

In conclusion, Adams’ work in Loserthink is to explain how sometimes truth and accuracy are at odds with practicality and effectiveness. There was no word to describe this process of thinking before, but Adams’ intent was to introduce exactly such a word into the language. He has succeeded with a few, but may take some more time to reach everybody.

Made to Stick by Chip Heath and Dan Heath

"They Laughed When I Wrote This Book Review. But When They Woke Up in an Ice-Filled Bathtub..."

1.

Can a book teach readers how to present ideas so that the ideas stick in other people's minds? Even readers who are not particularly creative? Even readers who can't deliver polished speeches? Even readers who don't have strong credibility or the resources of a PR firm?

You see, I have a problem.

I want to win the Star Slate Codex Book Review contest. But I expect a lot of other people are entering the contest, and some of them are probably way better at writing book reviews than I am.

I'm just a lurker at SSC. I've never posted a comment on the blog. I don't even read every post. I'm sure many of the regular readers and commentators are much better than I am at predicting what kind of book review will please you (Scott Alexander), and what kind of book review will get lots of votes from you all (SSC readers).

Heck, I'm not even sure I will be able to get you (Scott Alexander) to remember that my entry exists when you choose the finalists. And even if my entry becomes a finalist presented to you all (SSC readers), what I write may go in through your eyes and out your ears without ever reaching the parts of your brains which decide to vote for things. If I can't even pass the hurdle of 'you all can easily remember what the book review says when you make your decision' then my chances of winning this contest are zero.

This is where the book *Made to Stick* by Chip Heath and Dan Heath comes in. It's a book that I can review! It also promises to teach readers how to make their ideas stickier so that the audience remembers them when they make decisions.

Everything in this review is my attempt to put the techniques taught in the book into action. Do the techniques in this book really work? I really hope so. Otherwise, I'm not going to win anything in this contest.

2.

According to the book, one of the most common obstacles to getting ideas to stick is 'The Curse of Knowledge'.

What is 'The Curse of Knowledge'? (if you already know the answer, feel free to skip this part of the book review)

I'm going to demonstrate with an exercise where you try to remember sentences.

Example Sentence #1: 'Canada is in North America.'

Now, read that example sentence again, count to ten, and then, *without looking at the sentence again* say it out loud.

I bet most of you were able to pull that off. Maybe a few of you can't talk out loud or have some other reason why this was difficult for you, but for most of you, this is easy.

Example Sentence #2: 'Internal service funds don't use the flow of financial resources measurement focus.'

Now, I dare you to read sentence #2 again, count to ten, and then, *without looking at the sentence again* say it out loud.

I bet most of you failed.

Sure, a few of you may know what 'internal service funds' and 'the flow of financial resources measurement focus' are. You might even think that statement is just as obvious as 'Canada is in North America'. If so, you probably were able to do the exercise successfully, or at least say something like 'internal service funds don't use the current financial resources focus' which means practically the same thing.

But most of you are blissfully ignorant of what Example #2 means, which is why you have so much trouble keeping it in your head ten seconds after you read it.

Let's say we were at a meeting together, and in a sudden burst of passion I blurted out 'internal service funds don't use the flow of financial resources measurement focus.' Would you remember the content of what I said even a minute later? Unless you were cursed with the same knowledge as I am, the answer is no.

Now imagine that you were talking to someone who had never, ever heard of 'Canada' or 'North America' before. They have no idea that any place called 'America' exists. Heck, they haven't even heard of cardinal directions, so the concept of 'North' is completely new to them. Whereas almost all of you are cursed with knowledge of Canada and North America, this hypothetical person is blissfully ignorant of such things. How well do you think they would remember the sentence 'Canada is in North America' after counting to ten?

According to the writers of the book, you are overestimating your ability to communicate 'Canada is in North America' to people who are ignorant of 'Canada', 'America', and 'North'.

The book refers to a psychological experiment in which one group of people, called 'Tappers' tapped the rhythms of famous songs, such as "Happy Birthday to You" and "The Star-Spangled Banner" and a second group of people, called 'Listeners' tried to guess the song. Tappers predicted that the Listeners had 50 percent odds of guessing correctly. In reality, the Listeners guessed correctly only 2.5% of the time. That is because the Tappers could hear the melodies in their head, whereas the Listeners couldn't hear any melody.

If you are capable of mental imagery, you can probably visualize a map of North America and see Canada's position on the map, just as the Tappers can hear the tune of the song they are tapping to in their heads. But people are blissfully ignorant of 'Canada', 'America', and 'North' can't see that map in their minds.

The Curse of Knowledge is that people who know about Thing can't make good mental models of people who don't know about Thing. And then the people cursed with knowledge fail to communicate with the blissfully ignorant.

I wish I could have avoided writing this entire section of the book review and just use the phrase 'the Curse of Knowledge' whenever the concept came up. Some of you keep track of the names of cognitive biases and already knew what 'The Curse of Knowledge' meant. But if I relied on you all already knowing what 'The Curse of Knowledge' means, I would be falling into the 'Curse of Knowledge' trap myself. It's safer to assume that you need this explanation.

The Curse of Knowledge might cause me to lose this book review contest.

I am cursed with knowledge about *Made to Stick* because I read the whole book. Most of you haven't read the book, and thus are blissfully ignorant of its contents. I might throw around terms from the book like 'the Sinatra test' or 'antiauthority' or 'curiosity gap' which will just sail over your head.

If I fail to vigilantly counter the Curse of Knowledge, you won't understand what I'm trying to say in this review, which means you'll forget about it, which means I am going to do very badly in this contest.

3.

The book is built around the acronymn SUCCEsS, with each letter (except the final one) representing one of the principles of sticky ideas. They are also the titles of the chapters:

1. **S**implicity
2. **U**nexpectedness
3. **C**oncreteness
4. **C**redibility
5. **E**motions
6. **S**tories

I tried to cram every single one of these principles in the opening section of this book review.

SIMPLICITY - The main idea of this book review is "Can a book teach readers how to present ideas so that the ideas stick in other people's minds?" I made it the very first sentence so that you would not miss it.

UNEXPECTEDNESS - If this were the Kirkus Reviews, then the style of this review would be unexpected. Since this is Slate Star Codex, maybe this style isn't so unexpected. I also don't want to break the typical patterns for SSC book reviews so badly that I disqualify myself for the contest.

But I did try to create some 'curiosity gaps'.

According to the book, curiosity gaps are created when you give partial information that suggests missing information. If you have read the first half of a murder mystery novel, you probably know that at least one character is a murder victim. That suggests that there is also a murderer, but you (hopefully) don't know who they are yet. One piece of information (character

X has been murdered) suggests a missing piece of information (identity of the murderer). The mystery murder novelist is trying to create a compelling curiosity gap so that you will keep turning the pages.

I have told you that this book claims to teach techniques for making ideas sticky. What I have not told you is whether or not the techniques *work*. I'm also trying to make a curiosity gap out of how many votes this book review will get in the contest, but if this book review isn't a finalist that curiosity gap has already been closed for you.

CONCRETENESS

This book review contest is a concrete idea. Placing as a finalist or not is concrete. Numbers of votes cast are concrete. The book *Made to Stick* itself is a concrete idea, it's not some hypothetical abstraction of a type of book which may potentially exist, it's a specific book which already exists. I also tried to use concrete language such as 'go in through your eyes and out your ears without ever reaching the parts of your brains which decide to vote for things'.

CREDIBILITY

The book describes various different ways to establishing credibility. The one I chose is 'testable credential'. I'm using a bunch of techniques described in the book, YOU get to test whether they work or not. You don't have to take my word for it. Are you going to forget the contents of this book review ten minutes after you stop reading it, or are you still going to think about something I said here a week from now?

EMOTIONS

In retrospect, I could have put in an appeal to self-interest in the opening section, yet I did not. Maybe I thought that would have been too obvious. In case it's not too obvious, I'm going to add it here 'Imagine what you could do if you could make your ideas stick better in other people's heads - but make sure that techniques actually work before you invest time/money in this book.'

My first inclination was to address this book review to the third person - Scott Alexander and SSC readers. But the book strongly recommends using the word 'you' a lot, which is why I am addressing you all as 'you all'.

I also made an appeal to identity. I identified myself as a lurker. According to the book, that will get the attention of those of you who also identify as lurkers.

STORIES

The book recommends using one of three story patterns, not because they are the only story patterns which work, but because they are the three most common story patterns used for conveying sticky ideas.

I decided to go with the 'Challenge' story pattern. 'David and Goliath' is an archetypal 'Challenge' story.

In my story, the Slate Star Codex book review contest is the challenge. I am David. Goliath is the contestant who is submitting a brilliant review of a book about AI Safety, about the psychology of cognitive bias, about ancient Sumerian temple politics, about the replication

crisis, about Kaballah, about a brilliant topic which did not even occur to me. The book *Made to Stick* is my slingshot.

The book emphasizes that you don't have to invent your stories, it is just as good or even better to merely spot stories which serve your purpose. I didn't invent the story I'm using. I really am submitting this review to the book review contest. I really am daunted by the prospect of stiff competition. I really am nervous that you might laugh at me because this review is so gimmicky, or that you might sneer at me being so blatant about trying to not lose, or worst of all, that you might forget this book review ten minutes after you stop reading it.

4.

There is something I feel compelled to acknowledge, even though it's not directly relevant to the thesis 'does this book teach people how to make their ideas sticky?'

One of the examples featured prominently in Chapter 6 "Stories" is how Jared Fogle provided Subway with an extremely sticky and effective story which led to a dramatic increase in Subway's sales.

This book was published in 2007. Jared Fogle was convicted of distribution and receipt of child pornography and traveling to engage in illicit sexual conduct with a minor in 2015. There is audio evidence that he did much worse than that. As part of the plea bargain, he was required to pay \$100,000 each in compensation to ten different victims. But the writers of *Made to Stick* couldn't have known about any of that in 2007.

Though the writers' point about why the Fogle story was so effective for getting more people to buy Subway sandwiches still stands, I have a feeling that if they ever revised this book they would remove, or at least make substantial changes, to this example. Even from a detached amoral standpoint, Fogle's current reputation as a sexual predator of children is just too distracting to ever use him in an inspiring story again.

It's a reminder that something like 'convicted of sexually preying upon children' is very sticky, yet nobody wants to stick out in that particular way.

5.

I'm trying to think of examples which disprove the hypothesis 'this book teaches you how to make ideas sticky.'

I cannot think of any examples of ideas which stuck with me in spite of the absence of these techniques. Even when I was learning about things like internal service funds and the flow of financial resources measurement focus, I was filling in curiosity gaps - 'if internal service funds don't use the flow of financial resources measurement focus, what measurement focus do they use? The flow of economic resources measurement focus? But how are economic resources different from financial resources?'

So I have failed to think of an example of any idea sticking with me in spite of violating all of the principles in this book.

But what about the reverse? Are there examples of ideas which were presented to me in a way which perfectly executed the techniques in the book, yet they failed to stick with me anyway?

Well gee, I can't come up with any example like that either, because if I've forgotten something, *I've forgotten it.*

6.

A lot of the ideas in this book seem obvious, so obvious that I feel like I don't need a book to tell me about them.

Some of the advertisements on the SSC website are already using the principles described in the book, and I doubt it is because the people making those ads already read the book.

(Note: I wrote this before the SSC website got taken down: you can see the ads which I mention via [the Wayback Machine](#)

The book talks about why focusing on a single African child is so effective for fundraising for charities, and lo and behold, the Altruisto ad features a photo of a single African child. It's a cliché of charity fundraising, and thus non-surprising, but at least it's a tactic with a good track record of working to some extent (though, for the record, I've never clicked on the ad).

The 80,000 hours ad puts front and center a concrete idea - 80,000 hours. When I first started reading SSC, the 80,000 ad was the first one I ever clicked on because it forcefully presents a concrete idea. I had a curiosity gap about what '80,000 hours' had to do with doing maximum good with my career, and the idea of fulfilling career which maximises my contribution to solving the world's most pressing problems appealed very strongly to my enlightened self-interest.

"Giving What We Can" is a very simple idea, and that photo is an effective emotional appeal. When I see the group of happy people, I feel like joining the happy people and being happy alongside them. (That said, I've never clicked on the ad because I don't feel enough curiosity to learn more.)

The MealSquares ad even manages to use one of the more advanced techniques in the book, of building on the audience's existing mental schemas. Many SSC readers, especially the ones who would be most interested in MealSquares, already know what Soylent is, so "Think Soylent, except zero preparation, made with natural ingredients, and looks/tastes a lot like an ordinary scone" conveys a more vivid idea of what MealSquares are and what purpose they serve than if the ad had not mentioned 'Soylent' at all. (I clicked on this ad long before I wrote this review).

So if the techniques in the book are so obvious that everyone is already using them, what's the point of reading the book?

The point is that we often don't use the techniques in the book, even when they would benefit us.

I think Dr. Laura Baur is probably Cursed with Knowledge. She's an expert on 'relational psychotherapy' so she's lost sight of the fact that people like me don't know what 'relational psychotherapy' is. I don't even know enough about it to be curious. And isn't all psychiatry supposed to be 'rational'? And if not, how do I know that 'rational' psychiatry is better than 'irrational' psychiatry? The 'Seattle Anxiety Specialists' ad is better because I know what

'anxiety' is; I understand that they are the kind of people I might want approach for help if I have an anxiety problem.

I cannot recall ever paying attention to the Beeminder ad until I analyzed it for this book review. My reaction to 'Beeminder: Automated Self-Control Technology' is 'who cares.' I doubt I ever read the blurb before, because I think I would have at least remembered that it has something to do with keeping resolutions. Other than the phrase 'keep resolutions' it might as well be droning sounds to me, which is consistent with the bee theme but doesn't make me want to click on the ad. If instead, the ad had read 'Do you struggle with keeping resolutions?' (simple, emotional appeal to self-interest) I probably would at least have remembered the the ad exists, and possibly might have even clicked on it.

No, I'm not trying to shill for any of these advertisers, I'm only using them as my examples because because you can refer to them without leaving the SSC website.

Ironically, the techniques in the book may seem so obvious because it is much easier to remember sticky ideas (which use the techniques) than non-sticky ideas (which don't use the techniques).

7.

In Chapter 5 "Emotions" it is said:

Our publisher rejected the following subtitle for this book: "They Laughed When We Wrote This Book. But When They Woke Up in an Ice-Filled Bathtub..."

I decided to make that the title of this review because ~~I was too lazy to come up with my own title~~ I want you all to know what kind of taste Chip and Dan Heath have in (sub)titles.

8.

I am having doubts about the effectiveness of this book.

Presumably, the writers of the book have mastered the all of the principles and techniques in this book. Otherwise, they wouldn't be qualified to write a book about it, would they?

Presumably, the writers implemented all of the principles and techniques in the book itself. After all, if they know effective ways to get the content of their message to stick to the minds of their audience, they would want to use that in the book they wrote, right?

Presumably, because the writers applied super-duper sticky glue to the content of the book, the ideas are permanently stuck to my mind, right?

Ahem.

I do, in fact, remember much of the content of this book. But that might have something to do with me wanting to review it. I'm not sure how much would have stuck with me otherwise.

And in the middle ... I lost interest, and ended up reading another book instead. I couldn't come back to this book until I finished reading the other book. And no, the other book wasn't a murder

mystery, it wasn't even fiction, or a bestseller, it was a scholarly book for crying out loud, and I was *still* more eager to read it than *Made to Stick*.

Yes, I did come back to *Made to Stick* and read it to the end, but I'm not sure I would have done that if I hadn't already decided to write this review.

I think the introduction to *Made to Stick* is a little too successful at summarizing everything. Because the beginning of book had such a complete description of the main ideas, I only had mild curiosity about the main chapters. And indeed, the main chapters were not terribly surprising since the core ideas are already explained in the introduction.

By contrast, with the scholarly book, I discovered early on that I was bad at predicting the writers' opinions. I knew what the topic of each chapter was from the table of contents, but I didn't know what the analysis/judgment of the topic would be until I read the chapter. It was like binge-watching reaction videos on YouTube.

On the one hand, the *Made to Stick* framework can accurately describe why I was more motivated to read that scholarly book; it was more surprising and opened bigger curiosity gaps. So, *Made to Stick* earns points for having an explanation. On the other hand, this is an example of Chip and Dan Heath violating their own principle - they wrote a book that is too predictable. So I'm deducting points, and this book's score is back at zero.

If I'm David, and *Made to Stick* is my slingshot, I think I might be entering my showdown with Goliath with a broken slingshot.

9.

I just gave my very best try at implementing the principles of this book.

Now let's see how well the content of this book review sticks in your mind. If you have forgotten the what this book review says by tomorrow, that means the book didn't teach me how to make ideas stick in people's minds, and chances are it can't teach you either. But if a week from now you can still recall the gist of this book review, or a thought from this book review spontaneously floats into your mind, then maybe *Made to Stick* does live up to its promise.

On Numbers and Games by John H. Conway

John H. Conway recently died from COVID-19.
When Conway inevitably challenges the Devil to a game,
I would not envy the Devil.
When the Devil asks him what game he chooses,
Conway will simply answer:
"All of them. At once."

A Divine Game

Imagine you are God.
Literally.
Perfect in every way.
You create the universe,
naturally,
because this is what God is supposed to do.
You set the initial conditions perfectly,
because you are perfect in every way,
and let the rules of physics unfold from this initial condition
into a beautiful,
blossoming world.
This can take a while.
Since you created things perfectly,
there is nothing to do,
no need to intervene.
Things could get boring.
You need some way to entertain yourself.

So you let the Left arm play a game against the Right arm.
A Go game,
rich in majesty,
unimaginably big in scope.
Both the Left arm of God and the Right arm of God are,
of course,
great players.
They make no simple mistakes.

So after a while,
the middle of the board fills up.
The Divine Go game is now made up of many sub-regions
that do not interact.

When it is the Left arm's turn,
it picks a region,
and puts a stone.
When it is the Right arm's turn,
it does the same --
but not necessarily the *same* region.

A few billion years in,
God Himself looks at the Divine Go board.
"I wonder,"
He asks Himself,
"who is winning".
God wonders if it is possible to find some
"arithmetic of games",
so that if you analyzed each sub-region separately,
you would do some computation and figure out the state of the board.

Human Games

Since echoes of the Divine world are found in our world,
the same thing happened.
John H. Conway invented the Game of Life:
a 0-player game where initial conditions unfold into beautiful
guns shooting spacheships and,
potentially,
doing any computation imaginable.

Just like God,
when Conway was done creating a perfect world,
he decided to amuse himself with games between
Left and Right.
Just like God,
Conway looked for an arithmetic of games.

The first thing Conway did was separate the games that favor Left
from the games that favor Right.
He called the first kind "positive"
and the second kind "negative".

The games which favored neither he called "zero"
if the second player could win,
and "fuzzy" if the first player could win.

The reason that favoring the second player gets the nicer name
is because an empty game,
one with no legal moves,
favors the second player:
the first to play finds themselves with no legal moves,
and must concede defeat immediately.

Conway defined the "negative" of a game as one where the roles of
Left and Right are switched.
All of these,
any mortal human could do.
It was the next definition that secured Conway his Divine Right:
addition of games.

In order to add two games,
you imagine that at each point,
at each turn,
the player chooses one of those games,
and chooses a legal move.
Imagine adding chess and checkers,
and letting
Left begin.
Left likes checkers,
so they play a move on the checkers board.
Right prefers chess,
so they play a move on the chess board.
On Left's turn,
they could play again on the checkers board.
If you only looked at the checkers board,
you would see white move
twice:
clearly not a usual move in checkers.

If we know what addition is,
and what the negative is,
we can define subtraction: $A-B=A+(-B)$.
Now we can define an order on the games: A is less than or equal to B if A-B favors right or the
second player.

Numbers

Now Conway does something that every mathematician,
and even God Himself,
would envy,
He defines a number.
A number is a game where each Left option
(the game after Left took one turn)
is less than each Right option
(the game after Right took one turn),
and each option is itself a number.

So the empty game is a number
(conveniently called "zero"),
since there are no options for either player.
The game $\{0\}$ where Left can move to the empty game
but Right has no legal moves is clearly biased in favor of L:
it is positive,
and we call it 1.

Its negative, -1, is $\{0\}$.

The game $H=\{0|1\}$, where Left has the single option of moving to the empty game,
and Right has an option to move to 1,
which is a game they lose,
is clearly still biased in favor of Left.

But how biased?

We have defined addition and subtraction,
so with a little legwork, we can see that
 $H+H-1$ favors the second player:

in other words,

it is zero, and

H is exactly $1/2$.

It is still biased in favor of Left,

but in a strict sense,

less biased than 1.

It is possible to define multiplication on games,
but this definition is more subtle.

Conway himself struggled with the definition for many weeks.

But,

for the moment,

take it as a given that we can define multiplication,

and the expected results hold: $(1+1)$ times $1/2$ gives 1,

for example.

How many of the numbers we know are games?

"All of them" is a Huge understatement.

All real numbers are games.

So are all infinities.

You can embed all multi-variate real-coefficient polynomials in it.

They all can be added, subtracted, multiplied, divided
(except by zero, for this is a rule not even God or Conway can violate)
and compared.

Worlds

Conway built a machine to analyze the game of Go,
which splits naturally into separate sub-games where each player
chooses a sub-game and makes a move.

The "addition" operation is exactly the operation needed to analyze
such a game.

But what he got was a machine that
spits out a structure that takes a few weeks in an upper-level math class
to build,

all in the space of a few paragraphs.

Describing it as a math book does it injustice:

it is a physics book of a fictional universe...

or several fictional universes.

Together with No,

the field of numbers defined above,

Conway introduces us to the magical land of Oz,

and the mirror universe of On.

For much like God,

Conway wanted to create many worlds to maximize total happiness.

Open Borders: The Science and Ethics of Immigration by Bryan Caplan and Zach Weinersmith

Bryan Caplan is an economist at George Mason University and all-around interesting guy who is known for his out-there libertarian views about various social issues (especially education). *Open Borders: The Science and Ethics of Immigration* is his latest book, which argues for an end to all restrictions on migration and is in the format of a graphic novel illustrated by Zach Weinersmith (of SMBC fame). The first thing I would say about this book is that the graphic novel format works really well. The art style is cute and in general people seem to dismiss graphic novels because real reading is slightly boring, and therefore important. The next thing I would say is that the book makes an important argument on an issue where people have particularly poorly thought-out opinions. The data are pretty clear that immigration is massively more beneficial than most people realise – certainly economically, and perhaps socially too. However, upon reflection there are serious objections to open borders, and the argument of the book has a number of omissions.

The argument

Caplan really does believe that there should be no restrictions on immigration whatsoever, and that's exactly what his cartoon representation in this book argues for. The basic argument goes like this: people should, in general, be allowed to make decisions that they think will improve their lives, assuming they're not hurting anyone. Moving to a new country is exactly such a decision. Since immigrants often move in search of work, moving is associated with a massive increase in economic prosperity: by moving to the US and receiving *no* additional training or education, the average citizen of a developing country can expect their income to increase fivefold; for countries like Nigeria, the figure is tenfold. This is because developed countries are safer, more prosperous and have better quality institutions, so immigrants are more productive in them. The gains are so vast that a standard estimate is that open borders would *double* world GDP. And yet rich countries continue to restrict immigration, sometimes through formal caps on the number of entrants, and sometimes through complicated bureaucracy and paperwork which at best dissuades people from entering and at worst makes it literally impossible (like rejecting you for not filling out the middle name section on a form *even if you have no middle name*). Some of the arguments against immigration are xenophobic or racist, but many are legitimate concerns bought up in good faith. However, most (all?) of these are simply not borne out by the evidence. The consensus among economists is that immigration does not decrease natives' wages. Nor does it lead to an increase in poverty, crime, or a significant strain on the welfare state and social services. While the data about this is more unclear, immigrants seem to be barely different from natives in their political views and they adopt a lot of the cultural

values of their destination country. Hence, the contrary considerations are not enough to overwhelm our initial presumption in favour of allowing people to move and massively improve their standard of living, and so we should have open borders.

The objections

1: Parochialism

Open Borders is an extremely US-centric book. As someone from the land of 'not America', this is something that frustrates me about a lot of non-fiction. Caplan justifies his focus on the US by saying that his audience is mostly Americans and that that's where the highest quality data exists for. But in this case, the book makes a *way* narrower argument than is set out in the book's intro. This matters because, by focusing primarily on America, the case is made stronger than it otherwise would be. For instance, immigrants commit more crimes than native-born Europeans but fewer crimes than native-born Americans. Immigrants to the US assimilate unusually well (although some people say this is just because European countries are more regulated, and make such wise decisions as forbidding refugees from getting jobs - [link](#)).

The relentless focus on never talking about a rich country other than the US is doubly bizarre because the European Union has open borders between its member states! Surely analysing whether this has gone well should be the most convincing piece of evidence about open borders. Ireland is 17% foreign born, a significantly higher proportion than the US, and the immigration rate to Ireland has nearly *quadrupled* in the last 20 years. This would seem like a major success story of immigration. Meanwhile, Caplan only talks about the EU for a few panels toward the end of the book. Granted, there is more high-quality data about the US, and the audience is mostly American, but even still this doesn't seem to justify anywhere near the level of parochialism.

Until the 1920s, the US had de facto open borders, and this is another thing that I wish Caplan had dug into more. It certainly *seems* like America benefited a lot from immigration at this time (or, at the very least, that immigrant groups like the Irish did) but have people studied what the effects of open borders actually were? How did they change the quantity and quality of immigrants?

2: Humility

Open borders would be the largest social transformation possibly ever, and there isn't even that much research about it. We should in general be extremely humble about the prospects that our views about complex topics are completely right, and the downsides from open borders, if we are wrong, could be quite significant.

Caplan is unusually scrupulous at making sure his claims are backed up by the data. His book *The Case Against Education* is one of the most meticulously researched books I have ever read. So, it was a bit disappointing that there weren't more margins of

error attached to his claims. How confident are we that open borders would really double world GDP? 10%? 50%? 90%?

Natural experiments and case studies like this seem much more convincing to me. After reading a lot about the replication problems in economics and the colourful uses of statistics to get one's desired conclusion, I don't find projections along the lines of "open borders will double world GDP" very convincing.

3: Environment

Unless I'm mistaken, at no point does Caplan address the environmental harms of open borders. Moving people from low-emitting poor countries to high-emitting rich countries would lead to a pretty dramatic acceleration in global emissions. "Keep most of the world poor" is a terrible climate change strategy, yes, but there are some climate problems you might want to solve first before advocating for open borders. A world with open borders would be much richer, and so would have a lot more money to throw at the problem of climate change, but how much more would they throw? If the case for open borders were watertight, it would have to address this. I'm confident that Caplan has reflected and come to the conclusion that there are no climate problems that we can solve in a short-enough period of time to justify the harm caused by delaying open borders, but he doesn't show his work.

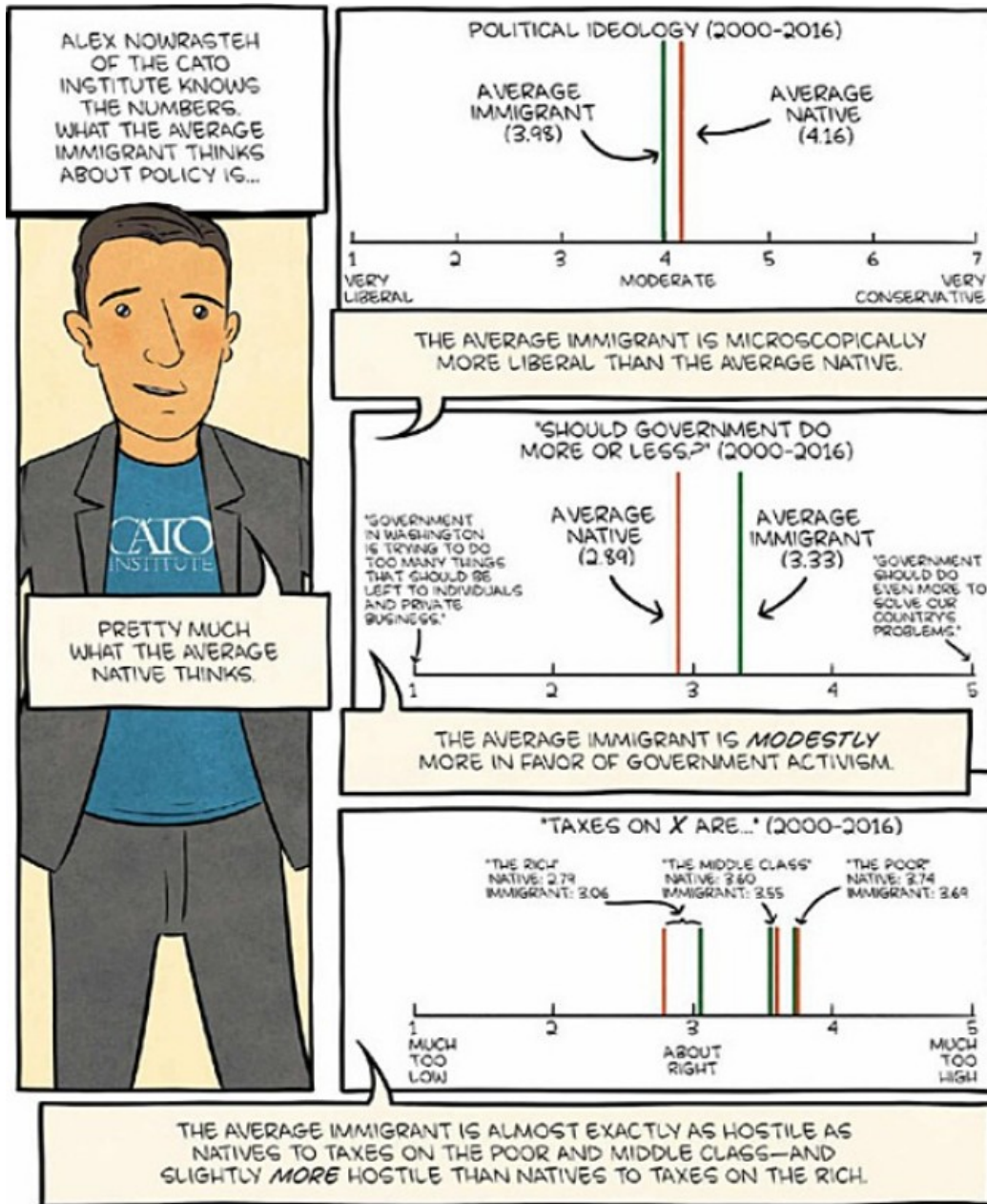
Sometimes, climate change gets used as an excuse for opposing almost any societal progress. This is unfortunate. But "open borders would create this gigantic problem, namely massively accelerated climate change, but the benefits outweigh the harms" was not the argument I got from the book. "Open borders are so good, and the objections are not that significant" was the argument I took away from the book.

There are considerations I can think of that would make the environmental objection less serious. Immigration would probably accelerate the trend of urbanisation, and cities are better for the environment (smaller houses, more use of public transport, etc.). People would also be able to move away from the regions that are worst affected.

Another serious objection is about the animal suffering that would be induced by open borders. I think that we should give a high degree of moral consideration to complex animals like cows and pigs, and that eating meat, 90% of which comes from factory farms, creates an almost unimaginable level of suffering. There are a couple of reasons why open borders would make this worse: the Western diet is more meat-heavy than diets from other places, and richer people in general consume more animal protein. Some people talk about the meat-eater problem: many interventions in global development look much less cost-effective if you give moral concern to animals (since, if the interventions save human lives or make people better off, they lead to greater meat consumption). The high levels of demand may further entrench factory farming as the default way meat is produced. This is not a consideration that most people have when thinking about open borders admittedly, but the premises are relatively uncontroversial. Virtually everyone agrees that animals are worthy of moral concern,

and many (most?) people see some problem with eating factory-farmed meat, even if they do not act on their discomfort.

4: Culture



Caplan has a section where he addresses the political effects of immigrants, largely drawing on data from Alex Nowrasteh at the Cato Institute finding that immigrants are a tiny bit more liberal than the general population but that their kids and grandkids regress

to the political mainstream. Immigrants and natives didn't have a partisan difference until the 1980s, and the partisan difference comes from immigrants being more likely to identify as independent, not from being more likely to identify as Democrat (although maybe after a while immigrants become acclimated and realise that third-parties never win...). This is interesting but doesn't address the tail risk of immigration leading to a Trump/Brexit dysfunctional level of polarisation or backlash (admittedly, that would be *very* speculative). It may be the case that the biggest harms from immigration come from people irrationally freaking out about immigration, but, uh, people are in fact irrational.

Here's Michael Huemer, in one of the most well-known philosophical defences of open borders, on the effects of immigration on culture - [link](#)):

Empirically, it is doubtful whether apprehensions about the demise of American culture are warranted. Around the world, American culture, and Western culture more generally, have shown a robustness that prompts more concern about the ability of other cultures to survive influence from the West than vice versa. For example, Coca-Cola now sells its products in over 200 countries around the world, with the average human being on Earth drinking 4.8 gallons of Coke per year. Footnote McDonald's operates more than 32,000 restaurants in over 100 countries.

This seems to kind of sidestep the objection. Mass migration to the US is not a concern because Coca-Cola will go out of business; it's a concern because democracy, freedom of speech, and the rights of women and homosexuals are deeply unpopular in much of the world. Importing millions of people from autocracies and societies that are otherwise deeply illiberal may well have adverse effects on democracy. This makes the case for having long waiting times for citizenship pretty good.

The selection effects right now for immigration to the US are really strong, but we have every reason to believe that this would decline under open borders. If immigration restrictions were lifted, the average quality of immigrant would almost certainly drop. This is something Caplan admits to, but the response to it didn't feel convincing. Just *how* much selection bias is there in who gets admitted and who doesn't?

5: Inequality

Caplan is an economist, so I can't confidently argue with his reasoning about the economics of immigration; however, I do have one gripe. The book is pretty convincing in arguing that immigration is the best tool we have for reducing poverty in an absolute sense, but what about poverty in a relative sense? Poor Americans still have it great by global standards, but they certainly don't feel that way, and the point of all this prosperity is presumably to make people subjectively better off. Don't get me wrong – I think that the average person *radically* overestimates how much of a problem relative poverty is. Defeating bona fide poverty – the type where people can't feed their kids –

seems to be priority number one, and a lot of the lack of appreciation of this comes from the fact that the large chunk of the world population whose lives are objectively terrible are not very visible to us. But still! Caplan is on the extreme end of this spectrum of only-caring-about-inequality-instrumentally.

Currently, the people who move from poor countries to rich countries are self-selected for being hard-working, intelligent, and conscientiousness. But what happens when the really unmotivated ne'er-do-well's start coming too? Under the current regime, these people would be relegated to the fringes of society. Could it even make some of them subjectively worse off, even if their pay cheque triples?

Caplan also doesn't really consider the extent to which racism and xenophobia might flare up in response to immigration (though he does have a great section covering the effects on social trust). The countries that are the closest to having open borders are the gulf states; they have many migrant workers from countries like Bangladesh and Sri Lanka. The economics side of my brain views this is great: the workers entered into a contract voluntarily, Qatar benefits by building cheap infrastructure, the Sri Lankans benefits by getting higher-paid jobs. But I do also fear this will lead to a kind of racially segregated dystopia.

In fact, immigrant groups would be largely stratified based on how wealthy they were to begin with. African immigrants would likely be deeply poor, followed by not-as-poor Indians, then richer Chinese, and so on. What happens to the politics and culture of a society that is that racially stratified? This is of course also a problem now, but I wonder what it might mean to scale it up so much. The fact that levels of education and training correlate with immigrants' ethnicity vis-à-vis the differences in wealth among countries would lead to a problematic level of statistical discrimination, at the very least. I worry about a circumstance in which society is so stratified, and labour market outcomes are so correlated with country of birth, that statistical discrimination is rampant, and gets confused with bona fide discrimination.

(Finally, regarding the welfare state, because I didn't know what section to put it in. One of the things that has dampened my enthusiasm for redistribution is that excess transfer payments aren't really compatible with high levels of immigration, and we know with a high degree of certainty that immigration is better at solving poverty than government programs. But does this actually happen? Do places that grow their welfare state subsequently shrink their level of immigration, or shift it toward higher-skilled immigrants? Is there something funky going on with cultural evolution such that support of immigration *and* welfare became tightly correlated beliefs?)

I initially was very sympathetic to the view – defended by some philosophers – that wealth inequality is not a problem per se; poverty is. But the more I think about it, the more this feels like squabbles over semantics. Yes, the *distribution* of resources is not intrinsically morally significant, but the mere fact that poor people don't have very much money isn't morally significant either. Conducting research about this is hard, and take the literature with a grain of salt, but, holding poverty constant, inequality seems to have lots of negative effects on all sorts of outcomes, including crime. So, given that it has negative outcomes, and is frequently caused by unjust social conditions, inequality is worth worrying about!

Conclusion

Toward the end of the book Caplan discusses whether it's a good idea to be advocating for open borders, or whether the idea is so radical that it will turn people off immigration even more. He comes to the conclusion that discussing open borders shifts the Overton window toward increasing immigration. I'm not so sure. For how important it is to convince people about things, I've seen remarkably little empirical research as to how you do it. Putting group polarisation aside, is it a good idea to give someone a stronger case or a weaker case to convince them to move their views in the direction of the argument? Regardless, the focus on economics and examining the realistic consequences of open borders

This book made me think about what low-hanging fruit might exist in the space of increasing immigration. As I mentioned, immigration in many countries is not formally capped but is de facto limited by being confusing and costly. Have people tried to start companies to fill this niche of streamlining immigration? Are there any foundations willing to run this kind of thing as a non-profit? Google turns up surprisingly few results.

All in all, I recommend this book. The thing it changed my mind the most about is the extent to which wealth is a function of where you are, not who you are. One estimate is that 60-70% of the global wealth disparity is explained by location alone. You could fix the institutions of poor countries from the ground up – but we don't know how to do this, it would take a long time, and it's unclear to what extent their problems arise from geography. Hence, the case for more immigration still looks pretty airtight. I'm excited to see the arguments for more radically reducing immigration restrictions be developed further!

Appendix: Other Arguments I Would Have Used

1. How much of India and China's economic growth is as a result of the fact that they're really big, and therefore, moving across them is a lot like immigrating? When Caplan pointed this out, I was pretty surprised I hadn't thought about it before. Were the two major economies to take drastic steps in reducing poverty in recent decades able to do so largely because they're really big? This is like one panel on one page, but I felt like he could have developed the argument more. In general, I think the book's argumentative style leans too highly on Estimates by Economists and not enough on case studies and natural experiments. Do more populous countries have greater growth in the long run? If so, this points us in the direction of open borders. Relatedly, I liked how Caplan talked about what Lant Pritchett calls 'zombie economies' – economies kept alive by restrictions that forbid people from leaving. A shockingly large share of the US has been declining in population for decades, yet we would regard it as absurd to say that people shouldn't be allowed to leave Nebraska because doing so would go against Nebraska's interests.

2. There are various arguments related to long-termism that Caplan didn't use; namely, the downsides of immigration (higher crime, perhaps draining the government's budget) are temporary but the upsides (higher economic growth) bear their fruit over centuries and will likely affect billions of future people. If you buy the argument, popular within effective altruism, that what matters most morally is our consequences on the long-run future, this would seem to be a point for open borders.

3. Caplan makes it seem like it's an open-and-shut case that immigration doesn't lead to an increase in unemployment. But many economists are also fans of the minimum wage. But surely there's a tension here? If the minimum wage has no disemployment effects, the labour market is perfectly inelastic, and if immigration has no disemployment effects, the labour market is perfectly elastic. So how elastic is the labour market?

4. Such a disproportionate amount of innovation comes from immigrants. More inventors immigrated to the US from 2000 to 2010 than to all other countries combined. Immigrants account for a quarter of total US invention and entrepreneurship. Maybe this is just because America disproportionately lets smart and innovative people move there. But maybe there are some agglomeration effects going on here specifically related to immigration? Immigration – or more particularly, clustering people together – seems to have been key to the success of various intellectual hubs throughout history, like the Bay Area recently, Vienna in the 20th century, and Edinburgh in the 18th century. This seems like a ripe topic for progress studies to tackle. Aesthetically, I agree with Caplan's choice not to talk about this much. People talking about all the “amazing contributions made by [insert immigrant group]” often comes off as condescending, in much the same way as token engagement with other cultures might. Make the case for immigration from prosperity and freedom, or don't make it at all! But it still has to be confessed that immigrants do seem to contribute a disproportionate amount – technologically, artistically, scientifically, and culturally – to the US.

5. I think there are good reasons to believe that way fewer immigrants would actually move than Caplan presupposes. During the entire Greek financial crisis, only 3% of the Greek population moved country (!), at a time when the unemployment rate was 27% - and remember, Greek have more than a dozen prosperous destination countries to choose from *with no paperwork involved!* Inertia is the most powerful force in the universe. Caplan's defence of his high implicit estimates is that, once the ball gets rolling, more and more immigrants from a particular country will move – for instance, immigration from Puerto Rico to the US was lower than you would assume given the difference in economic opportunity, but then Puerto Rican communities formed in many US cities, and more and more people moved. Gallop found that more than 100 million people want to migrate to the US. 750 million say that they would leave their home country if they could. But we have reason to doubt people would actually act on this. This makes open borders a little more palatable to people that are sceptical of immigration: it wouldn't be as different to the status quo as you might expect.

6. What factors have led Canada and Australia to handle immigration so well?

7. There is a general perception that Muslim immigration to the EU has gone poorly. How much of this is hysteria? Why would future rounds of immigration not have problems in the same way?

8. Greying is not something that Caplan talked about much. At first this might seem surprising – one common-sense case for immigration is that people in Europe and America are getting too old to work and they need immigrants to replenish their workforce.

9. There aren't really jobs that "Americans won't do", since, if people don't like doing something, the wages will rise until they start doing it to meet demand. However, this price is such that there's significant deadweight loss – mutually beneficial trades that no longer occur. For instance, more people would get more childcare if the government allowed more immigration. Caplan discusses this, but I didn't feel sufficiently inspired to think about how this would be great for me personally. If I lived in a place with open borders, I'd probably hire a personal assistant or something.

10. Another book I read recently and recommend is Matt Yglesias' *One Billion Americans*. He argues for large-scale population growth, partially through immigration but mostly through an increase in fertility, to maintain American pre-eminence over China and India. He argues that, for all of its failings, American dominance is better than the alternative. And America is at a disadvantage on this front by having a billion fewer people than the Asian giants. I'm not sure this argument should have gone in the book – it would take a long time to justify, and open borders appeals to a lot of left-libertarian sensibilities that might be offended at the idea of American global hegemony. But it would be an interesting project for the open borders community to look at the geopolitics of population growth. How important are marginal increases in population to geopolitical power? Are spurts in population growth followed by increases in various measures of hard power? Soft power?

Power, Sex, Suicide: Mitochondria and the meaning of life by Nick Lane

Many works try to take a weirdly shaped lens through which to rechart a small portion of our map of the world. [Power, Sex, Suicide](#) does that through the lens of, ahem, mitochondria. The weirdest thing about it is how splendidly it succeeds.

I - What is this book about?

Almost 2 years after originally reading it, this beautiful aberration has remained ingrained in my brain in ways that surprise me.

This book is many things, among them, it's a reply to Andy Matuschak's and others that don't believe in a book's power to teach you technical topics.

Nick Lane somehow managed to pack an introduction into mitochondrial science together with an overview of the most fundamental questions of biology and a fair dose of storytelling and experiment-describing into a dry coat of British humor.

How did life originate? Why do we die? Why do we have sex and sexes? Why do we eat? How do we differentiate into species? How do paradigms and theories in the field of biology shift? How does evolution work? What causes cancer? How did the immune system evolve ?

Are among the questions addressed by this book. They are all questions I had decent answers to. But somehow this book managed to make them all seem new, gave it's own somewhat odd but seemingly very correct answer, and left my brain with a slight but observable perspective tilt in all of those areas.

It's hard to think of a person this book is aimed at specifically. It seems too technical for a lay audience and explains too many basic concepts to be aimed at people working in biochemistry, molecular bio, or cell bio. Yet it's hard for me to think of a single intelligent person to which I wouldn't recommend it.

II - What is this book actually about?

This book is about mitochondria, about trying to give a detailed and awe-invoking picture of this organism-organelle that's so foundational to eukaryote life.

It outlines theories about the bacterial progenitors of mitochondria and how the original symbiosis might have come about. From there it outlines things such as whether or not it happened multiple times and how in the world it could have happened only once.

It contains an engaging historical telling of how scientists discovered biological energy creation, it explains how mitochondria function, how previous generations thought they did, and how future generations might correct our misconceptions.

It grapples with why mitochondria exist in the first place, why they seem to be so crucial to tissue-differentiate organisms, why their genetic material ought to be preserved independently from that of the cell.

It does this in such a holistic way that it brings tears to my eyes. Presenting both the theoretical and experimental reasoning chains and trying to be so even-handed with each hypothesis that it feels hard to know the author's favorite, even though he outright tells you which one it is.

It spends 12 pages to reach a short "correct" statement like:

The speed of respiration is very sensitive to changing circumstances—whether we're awake or asleep, or doing aerobics, sitting around, writing books, or chasing a ball. These sudden shifts demand that mitochondria adapt their activity at a molecular level—a requirement that is too important, and abruptly swinging, to be controlled at a distance by the bureaucratic confederation of genes far away in the nucleus.

But the core of this book is not formed of correct statements, but rather in the impressively large and well-argued inquiry chains, the author uses to write them.

I remember that Scott's review of [The Structure Of Scientific Revolutions](#) laments:

one of my big complaints about this book is that, for a purported description of How Science Everywhere Is Always Practiced, it really just gives five examples

This book contains at least 4 examples that could easily amend Khun's work and map it to biology. I found the most engaging one to be the foray into how eukaryotes produce energy.

It's a slowly progressing dance between current knowledge, available assays, and the theories resulting from the two. In nearly 50 pages the author goes from a picture of energy creation focused around the idea of fermentation, inspired by wine-making, to our modern conception of mitochondria with their fancy membrane including the 4 types of respiratory complexes chartering through various compounds to maintain the delicate balance involved in not killing our cells due to lack of ATP, while also not killing our cells due to overwhelming them with ROS or depleting them of NAD+.

Even though the book is about mitochondria, how they work and how we can view the world once we understand that, it's just as much a book about the living-breathing scientific method as it was applied throughout the ages

III - But now, really, what is this book actually about?

Ok, this book is probably not actually about the evolution of science. The author uses the "ask a question, look at the progress of answers through history" approach, but I think many books do that.

This book is about the origination of eukaryote life. After 70 pages of prefaces and introductions, the author's ponderings around the origination of eukaryotes open the first real chapter:

The void between bacterial and eukaryotic cells is greater

than any other in biology. Even if we begrudgingly accept bacterial colonies as true multicellular organisms, they never got beyond a very basic level of organization. This is hardly for lack of time or opportunity—bacteria dominated the world for two billion years, have colonized all thinkable environments and more than a few unthinkable ones, and in terms of biomass still outweigh all multicellular life put together. Yet for some reason, bacteria never evolved into the kind of multicellular organism that a man on the street might recognize. In contrast, the eukaryotic cell appeared much later (according to the main-stream view) and in the space of just a few hundred million years—a fraction of the time available to bacteria—gave rise to the great fountain of life we see all around us.

Solutions are suggested, though not always clearly delimited, sometimes it feels like speculation is just being thrown at the wall and blending into each other to form what could be the basis of hundreds of slightly different valid hypotheses.

We are presented with Nick Lane's favorite theory on the subject, the Hydrogen hypothesis, but after reading the chapter about it twice it still feels like a mess cobbled out of the part of different

theories and observation rather than a coherent whole. This makes me hopeful that Nick Lane is actually presenting the ideas as they are rather than taking a popsci hat and trying to do the whole "Unification of the field via glancing over every single interesting question using a metaphor or disclaimer".

The hydrogen hypothesis can be summed up as:

While methanogens are uniquely resourceful in their metabolic powers, they nonetheless face a serious obstacle... . The trouble is that, while carbon dioxide is plentiful, hydrogen is hard to come by in any environment containing oxygen, as hydrogen and oxygen react together to form water. From the point of view of a methanogen, then, anything that provides a little hydrogen is a blessing.

Hydrogenosomes are a double boon, because they release both hydrogen gas and carbon dioxide, the very substances that methanogens crave, in the process of generating their own energy. Even more importantly, they don't need oxygen to do this—quite the contrary, they prefer to avoid oxygen—and so they function in the very low-oxygen conditions required by methanogens. No wonder the methanogens suckle up to hydrogenosomes like greedy piglets! The insight of Martin and Müller was to appreciate that this kind of intimate metabolic union might have been the basis of the original eukaryotic merger.

However, the books don't leave it at that and pick up the thread all the way to complex eukaryotic lifeforms. In the process, some interesting theories that don't seem related to mitochondria are approached. The one that stuck with me is how our immune systems may well have developed from the processes bacteria used to signal others in their colony when they wanted to get some DNA transferred to them. Since this usually happened when a bacteria was under some form of inter-cellular stress. This signal persisted in eukaryotes to form the basis of the much more complex interactions between cells and our immune system. How mitochondria are involved in all of this I leave to the reader to discover, however, as, with everything in these books, this thread of ideas does get tied back into them.

IV - So, is that all?

I keep feeling like I'm doing a disservice to the book by presenting its ideas here because I wouldn't want to tie its content to just those few points. This is a work to be enjoyed one chapter at a time, you could probably go in mid-way through and you'd still enjoy it, you could pause in the middle and read the rest 1 year later, and you'd be right back on track. So honestly, I can

only describe bits of certain chapters, because to describe the whole would be to rewrite the whole book.

There is one thing though, which this book made me realize, which only clicks when it's read as an individual unity. Which is a glimpse into the thought process of a specialist in a field.

I don't mean a glimpse into his knowledge of his particular field, but rather a glimpse into how he integrates "the rest of science" in a mindset where it becomes subservient to his niche interest, in a way that enhances it, that gives the rest of us a new perspective on that "rest of science".

Even more so, this is done in a very respectful and, ahem, scientific? manner. In that, the author is never cherry-picking suitable evidence from other fields. More so, he seems to be doing in-depth dives and bringing in both the things that confirm his views and those that threaten them. Indeed, to the author's credit, I'm not even sure what his personal views are on many of the contested subjects approached by the book, I might even go as far as to compliment him in stating that he has none.

What he does have, however, is a strong fascination with creating edges in the graph of human knowledge between mitochondria, energy production, eukaryote evolution, and everything else. At least that is the feeling I get from reading the book. I find this kind of healthy obsession to be rare and fascinating, very enjoyable to observe, and to voyeuristically partake in for a few hours.

In my library of pop-biology, this book stands in a place of honor, next to Sapolsky's course in human behavioral biology and *The Selfish Gene*. I can't help but recommend [it](#).

Probability Theory: The Logic of Science by E. T. Jaynes

$P(A|B) = [P(A) * P(B|A)] / P(B)$, all the rest is commentary.

- Astral Codex Ten tagline.

If I told you that over the course of last year I have read Hartshorne's graduate math textbook "[Algebraic Geometry](#)" and did all the exercises in it, you would probably assume I have learned some mathematics.

If I told you that over the course of last year I have read Thich Nhất Hạnh's "[The Miracle of Mindfulness](#)", and practiced all the techniques in it, you could deduce that I have developed some applicable skills stemming from a metaphysical doctrine.

In reality I did neither of these things.

I did, however, read E. T. Jaynes's "[Probability Theory: The Logic of Science](#)" (PT:TLoS from here on) and [solved \(most of\) the exercises](#).

If you [have heard anything](#) about this book, you may have expected that I have learned some mathematics, and developed some applicable skills stemming from a metaphysical doctrine. In reality, I have learned that in the 20th century disputes concerning probability and statistics, physicists whose last names start with J were (almost) always right, and everyone else was almost always wrong. Ok, fine, I did learn some math, though mostly in pursuit of understanding of an offhand remark or a solution to an exercise, often by following the crumb trail of hints and looking up references left by Jaynes in the text, and only occasionally from the text itself. As for the metaphysical indoctrination, well, there was a fair bit of that - but one does not simply join the [Bayesian Conspiracy](#) by reading a 700+ page book. One must read a [1600+ page book](#) at least!

On the origin of PT:TLoS.

Edwin Thompson (i.e. E.T.) Jaynes was a Ph.D. student of Eugene Wigner, the [unreasonably effective](#) Nobel laureate physicist. Wigner is reported to have later characterized Jaynes as "one

of the two most under-appreciated people in physics." Jaynes's PhD thesis was on ferroelectricity, and apart from contributions to probability and statistical mechanics, he is perhaps most known for his [work](#) in quantum optics.

Jaynes defended his Ph.D. at Princeton in 1950, and then moved to Stanford. He did what one is supposed to do there: invested in a [Palo Alto tech startup](#). Since his first field of research could be called applied classical electrodynamics, he also consulted for the startup, calculating behavior of electrons in cavity resonators and working on magnetic resonance. Apparently this led to him buying a fairly large house -- though this was the 1950s, when [normal people](#) could have houses in Palo Alto. He bought an even larger house when he moved to the University of Washington at St. Louis in 1960.

In 1957 Jaynes published [two papers](#) on "Information Theory and Statistical Mechanics" concerned with formulating (the Gibbs's picture of) statistical mechanics in terms of information theory, first for classical and second for quantum systems. At about the same time he delivered a series of lectures on "Probability Theory in Science and Engineering" at the Field Research Laboratory of the Mobil oil company. The [published version](#) of 5 of these lectures is the first draft of the PT:TLoS. It includes a now-extinct section on the Gibbs model and one titled "why does statistical mechanics work?", as well as (much) briefer versions of chapters 1, 2, 4, 5, 6, 11, and 18 of PT:TLoS, for a total of about 200 typed pages overall. It also contains a "historical introduction" explaining "how it could happen that a person who is a rather strange mixture of two thirds theoretical physicist and one-third electrical engineer could ~~grow up to be a hero and a scholar~~ get really worried about the foundations of probability theory". The answer, of course, is by "trying to understand what statistical mechanics is all about and how it is related to communication theory". I'd say that it's a struggle that still goes on for many of us!

Jaynes says that "in the years 1957–1970 the lectures were repeated, with steadily increasing content, at many other universities and research laboratories." In 1974 some of this steadily increasing content was assembled into a 446 page "fragmentary edition" entitled "Probability Theory With Applications in Science and Engineering" with a stated goal of eventually having "approximately 30 Lectures" in the project. It now also included some of what will become chapters 10, 13, 19, and 22 of PT:TLoS, as well as a chapter on irreversible statistical mechanics.

Jaynes continued working on this material up until his retirement, and even more so after it. The magnum opus was woefully unfinished at the time of Jaynes's death in 1998. By 2004, its manuscript was shaped into a book by Jaynes's former graduate student Larry Bretthorst, resulting in the 727-page commentary on Bayes's theorem that we are now reviewing.

What Jaynes taught.

While no [Australian fashion models](#) seem to be available to distill the core idea of PT:TLoS into a single passage, we can get something reasonably close from Jaynes himself. Right from the start, he declares: "Our topic is the optimal processing of incomplete information", and the focus is on producing "quantitative rules for conducting inference". Note that while other frameworks might "process incomplete information" by [learning hypothesis consistent with data](#), Jaynes is after not just good enough processing, but the "optimal" one. Of course, the "quantitative rules" mentioned turn out to be those of "probability theory and all of its conventional mathematics, but now viewed in a wider context than that of the standard textbooks." This is the essential content of theorems of Richard Cox and Jaynes spends the first chapter fleshing out more precisely what the "quantitative rules for conducting inference" are and how they should look like. The second chapter is spent reproving Cox's results (i.e. that only probabilities allow us to do inference the way we would like).

With this first (but by no means last) tussle with foundations out of the way, Jaynes proceeds to develop some of the math needed for basic applications in 'direct' and 'inverse' probability. Here, by 'basic applications' I mean counting balls in urns (lest you find this boring, let me remind you that counting things in urns, is not only a centuries-old pastime of probability theorists, but is [essential for the functioning of any democratic society](#)). By 'direct probability' I mean things like: if there are a hundred red and a hundred blue ~~balls~~ balls in an urn and you draw 10 'at random', what is the probability that they are all red? That 9 of them are red and 1 is blue? Et cetera. This is 'sampling theory' and is covered in chapter 3, with the question of what 'at random' means getting some love in section 3.8.1. 'Inverse probability', on the other hand, is the old-school name for the more interesting kind of question: suppose you draw 10 balls at random from an urn containing 200 balls, and all 10 are red (this is your 'data'). How likely is it that there were 0 red balls in the urn? How about 1 red ball? How about 100? Here of course the answer depends on what we thought about the number of red balls in the urn before doing the drawing -- if we have looked in the urn just before and counted the balls directly, the drawing itself is unlikely to change our opinion about this 'prior' count. This **prior** is the remaining necessary ingredient. Once we have it, Bayes's theorem - the $P(A|B) = [P(A)*P(B|A)]/P(B)$ from the tagline - finishes the job. This is the famed 'Bayesian update' producing the new, i.e. posterior, beliefs from the prior ones and the probability of data. The innocuous-looking observation that inference requires priors is the fact that launched a thousand ships. Jaynes lists 4 "principles" for obtaining this missing ingredient (you know it's bad when there is more

than one "principle", and more than two is real trouble). He postpones further discussion to later chapters and proceeds to develop 'inverse probability' - aka hypothesis testing - assuming the prior is known somehow. Along the way in chapter 4.2 we get introduced to measuring information (or 'evidence') provided by the data in decibels (which I believe Jaynes invented independently of the equivalent "[decibans](#)" of Turing and Good), and learn how to do multiple hypothesis testing in chapter ~~86~~ 4.4.

With all this hard work out of the way, we get to "queer uses of probability theory" ~~also known as the seeds of the CFAR curriculum~~. While non-technical, this chapter explains how to reason "in a Bayesian way" about telepathy, why the same evidence presented to different people may make their opinions diverge more, how the Bayesian nature of visual perception may explain optical illusions how not to weigh evidence in court, and other useful things like that. "It's the priors, stupid"- for the most part; yet the details are entertaining and sometimes illuminating.

By chapter 6 the break is over, and we return to our urns. Amid some rather mundane calculations, some inspiring things happen. Under the rubric of "effects of qualitative prior information" (here 'qualitative' means something like knowing 'who does what to whom'), Jaynes introduces what we now can recognize as rudimentary probabilistic graphical models. The question of the choice of a prior returns briefly, only to be postponed again. For the most part it is a continuation of what has gone on before.

Chapter 7, dedicated to the Gaussian distribution, is a change of pace. While mathematically interesting, at first blush it may seem purely technical. Yet there is a key question behind it: why is Gaussian distribution so ubiquitous? Of course, mathematical reality being what it is, all good explanations are connected to each other; but the side from which one approaches the network of explanations matters both philosophically, and in terms of what further ideas it generates. Here, as in many other situations, Jaynes has a favorite side.

A standard answer is commonly taught: if a number we are considering is a sum of many (sufficiently) independent random pieces the result will be approximately Gaussian. Since many things have multiple 'small causes', this is a common situation. Mathematically, this is expressed as the Central limit theorem. A mechanism that makes this work also explains why Gaussian distribution is connected to the least squares fitting of linear models, and, more generally, illuminates why mean and variance are the only thing that matter in a Gaussian distribution. Thus Jaynes's favorite explanation is reached: Gaussian distribution is the one we would obtain if we agree that we know some random number's mean and variance, and nothing else. For example, we might be analyzing a noisy current in an electric circuit, and we know that the noise is zero on average, and also know its average power (which gives us its variance), but

we don't really know much else about this noise. Then, as Jaynes promises to show later, the Gaussian is the distribution of **maximum entropy** subject to our knowledge, the one expressing total ignorance beyond those two values. Thus, out of a technical sounding-question in a technical-looking chapter a major theme is born: if you know something, and want to get a prior reflecting that knowledge and nothing else, look for a maximum entropy distribution compatible with this knowledge. This maximum entropy principle is one of the four principles for finding priors that Jaynes mentioned back in chapter 4, and Jaynes is widely known for advocating it. Jaynes was "a perennial participant" of the annual workshop on Maximum Entropy and Bayesian Methods which was run from 1981 and until his retirement. [Several volumes](#) of the workshop's proceedings are [dedicated](#) to Jaynes. Properties of maximum entropy distributions (at least for 'finite' situations) are explored in chapter 11. This is also where the editing seams start to show: producing Gaussian distribution as a maximum entropy one is easy after the material in chapter 11 is absorbed, but as far as I can tell Jaynes never actually gets back to fulfilling his promise to do so, and does not even assign this as an exercise to the reader. Chapter 11 is in part II, where completeness of the text begins to decline.

Another one of those four principles for finding priors is "group invariance" (more properly "equivariance"). It is explored in Chapter 12 of PT:TLoS. The name hides a simple idea and a surprising complication. Here is the idea: if your setting is unchanged by some modification - and this includes your state of knowledge - then your prior should be unchanged by this modification. For example, if I don't know anything about the length of something then I don't know anything about twice its length. The modification here is stretching things by a factor of 2. Now, if I think my ignorance about these two situations should be expressed the same way mathematically - then my prior should be unchanged by the stretching. It turns out that in many situations this suffices to mostly determine the prior. In the above example of length (technically known as 'a scale parameter') I conclude that my prior needs to be such that, for any L , the probability that the length lies between L and $2L$ is the same as that it lies between $2L$ and $4L$, is the same as that it lies between $4L$ and $8L$ etc. I then must conclude that the resulting prior probability density at length x is proportional to $1/x$ (indeed, observe: the integral of $1/x$ between L and $2L$ is the same as between $2L$ and $4L$ and so on).

The surprising complication is that often this is not enough. For the simplest examples - like the 'scale' one above - this complication does not arise, but for the case of determining 'scale' and 'location' simultaneously it already does, and Jaynes gets it wrong. A deeper analysis of the situation hinges on the difference between something called 'right-invariant (Haar) measure' and 'left-invariant (Haar) measure'. A [book](#) of James Berger explains that the "correct" one to use is the right one. Jaynes certainly knows about this book, since he refers to it several times elsewhere in PT:TLoS. In his generally very positive and friendly [review](#), Stanford statistician

Persi Diaconis mentions that Jaynes has been accused of "not knowing his left from his right Haar measure". In fact, in PT:TLoS Jaynes seems wholly oblivious to the issue in the first place. His language is sufficiently imprecise to be confusing rather than enlightening -- which is doubly strange since the explanations in Berger's book are considerably clearer. I should note that in the 1974 "fragmentary edition" one can find a rather genteel "word of explanation and apology to mathematicians who may happen on this book not written for them", excusing the absence of measure-theoretic notions. Jaynes says: "I am not opposed to these things, and will gladly use and teach them as soon as I find one specific real application where they are needed." In the PT:TLoS the rejection of modern mathematical toolkit continues unabated, but any tone of apology is gone. Perhaps, at least in Chapter 12, some measure theory could have been useful after all.

But enough of that. All of this "inference" business is about what to think, and who cares about that. We want to know what to do! Thus, we need decision theory. The shift in focus from inference to decision gives an occasion for some discoursing on British vs. American priorities in life. This is particularly amusing given that the main credit for decision theory goes to the [Hungarian](#) mathematician [Abraham Wald](#), of the "[it's the missing bullet hole locations that you need to worry about](#)" fame. (Wald's dramatic life story is second perhaps only to that of [Alexander Grothendieck](#) in its Hollywood potential.) Wald's decision theory proceeds by assigning to each possible action (say: buy, sell) some utility, dependent on the true state of the world (say, the price tomorrow). The recommended action is then the one that maximizes the expected utility, 'expected' meaning averaged over your beliefs about the true state of the world (i.e. today's beliefs about tomorrow's price). That is, ignoring transaction costs: buy if the expected utility of tomorrow's price is higher than the utility of today's price, and sell otherwise. (Of course the economists being naturally dismal talk about minimizing loss - or cost - rather than maximizing utility.)

This may sound trivial, but that's because, under Jaynes's influence, we are already talking in the language of **beliefs about the true state of the world** -- what a statistician may call distribution of the model parameter, something which is not really allowed in the 'orthodox' or 'frequentist' approach to statistics. Instead, a frequentist might be concerned with a 'decision procedure' or 'strategy' based on some data, i.e. some process that takes in data and spits out the action to take. This procedure should not be too wild, and what "not too wild" means is formalized by Wald and is given the name "admissible". (Jaynes seems to interpret 'admissible' as 'good' and proceeds to rally against this term, by providing some not-so-good admissible strategies; I think simply interpreting 'admissible' as 'not obviously stupid' would've ameliorated that particular pet peeve.) This sets up a triumph of Bayesianism: many years after starting the study of admissible strategies, Wald proved that they are all basically Bayesian. That is, they

are equivalent to starting with some prior 'beliefs about the true state of the world', updating them based on the data - via Bayes's theorem, of course - and then choosing the action that maximizes expected utility. Moreover, in the case where the "decision" is actually estimating a parameter, by varying your utility/loss function and applying the above strategy, you may recover some standard estimators, such as taking the posterior mean, or taking the maximum of the posterior, of which classical maximum likelihood is a special case. Jaynes rightly points out that the shape of the loss function can change the decision quite drastically: in deciding between cutting your hair too short or too long, one type of error is much less costly than another; the cost of various errors in a 'William Tell-type scenario' is even further from the usual models.

With this - essentially final - layer of theory, we are ready for some applications. One application Jaynes considers is distinguishing the signal from the noise. And he does mean signal - an electrical one, in volts (it is probably that "one-third engineer" in Jaynes speaking). Another application is deciding what widgets to produce in an imaginary widget factory. While the first, simpler, task is arguably [more important](#), it is the latter that is more revealing of both Jaynes's process and its flaws. The analysis is fine - great even - when taken on its own, but there are no sanity checks and no robustness analysis. If I actually had a widget factory, I would not act based on this whole thing, at least not before hiring someone to vary the model and see how it flexes.

An important issue that remains is how to summarize beliefs about the world. Imagine I have a coin. I may say that the probability that it will land heads on the next toss is half, but this does not capture all my beliefs about the coin. Perhaps I have personally forged the coin to be fair, or perhaps I have never seen it before in my life. Now, imagine I see it be tossed and come up heads 10 times in a row. What would be my prediction of the next toss now? In the first case it is still pretty close to 50-50 ('pretty close' rather than 'exactly' because maybe my manufacturing process was flawed, or maybe the throws were rigged, or maybe I'm insane - all good reasons for me to hedge against [being overconfident](#)). In the second case I might start to suspect that the coin is not fair, and adjust my forecast accordingly. The question before us is how to account for this difference. Jaynes takes this up in chapter 18, and essentially invents a two-level [hierarchical Bayesian model](#). It goes roughly like this: First, for each p , I consider the world in which the coin is biased to land heads with probability p , and judge how likely it is that I am in that world. I collect all these judgments into what Jaynes calls "the A_p distribution", A_p being his notation for the statement 'the coin is biased to land heads with probability p '. Then, once the results of the coin flips come in, I update this A_p distribution using Bayes's theorem. The difference between the two scenarios above ('personally forged' vs 'never seen before') is in the initial shape of the distribution for A_p . The 'I forged this coin' initial distribution has a high peak

near $p=0.5$, while the 'this is just some coin' one is more spread out. They both average to 0.5, which is why they both, before any data comes in, lead me to predict 50-50 for the next toss. However, the first one is a more confident prior and is less susceptible to change based on new evidence. (Incidentally, if our initial distribution A_p is in the [Beta family](#), then updating it is particularly [easy to do](#), which is what makes the section 18.5 work out). One thing to note here is that we are now talking about something like 'probabilities of probabilities', and this is not what Jaynes discussed when setting up the whole "extension of logic" business. In fact, I agree with the [the contention](#) that 'logic' in 'the logic of science' is to be, at least initially, understood as propositional calculus, since this is actually what Cox's theorems deal with. Finding probabilistic extensions of predicate (and higher-order) logic seems to be the subject of some current research. Whether this has some bearing on Bayesianism as "a complete theory of formal rationality" is a question slightly too philosophical for my usual tastes.

All of this is no doubt very thrilling: I mean, we are "only" solving the question of how one should reason - and act! - in the world. We call it 'inference' just to keep the excitement down and keep philosopher-logicians off our back. But it is not nearly as much fun as the numerous polemical tirades against "the orthodoxy", be it of Fisher, Pearson, or Feller patriarchate.

À la recherche du temps perdu.

Chapters 8, 16, and 17 give some account of - and Jaynesian commentary on - the classical statistics. These did not appear in the earlier drafts, which were more focused on expounding Jaynes's own theories. In PT:TLoS, Jaynes spends quite a bit of time in remembrance of things past, recounting his disagreements with classical statisticians. Their "orthodoxy" is described in terms of its "pathology" and "folly". Jaynes's main charge is that their methods are "ad hoc" - a phrase that appears 47 times in PT:TLoS. For the book whose chief aim is to develop systematic rules of inference, this is probably not surprising.

If one were to pick out a single antagonist in the PT:TLoS it would have to be Sir Ronald Aylmer Fisher. One could say that Fisher was a geneticist and a statistician. Or, one [could say](#) that he was "the greatest of Darwin's successors" and "the single most important figure in 20th century statistics". Bradley Efron (another Stanford statistician) [writes](#) that "one difficulty in assessing the importance of Fisherian statistics is that it's hard to say just what it is. Fisher had an amazing number of important ideas and some of them, like randomization inference and conditionality, are contradictory. It's a little as if in economics Marx, Adam Smith and Keynes turned out to be the same person."

Among many charges Jaynes lays at Fisher is that of establishing statistics as a collection of (ad hoc!) recipes for analyzing data. In Jaynes's view Fisher's cookbooks (primarily "[Statistical Methods for Research Workers](#)", but also [The Design of Experiments](#).) established the situation in which a scientist was to follow the recipes, but was not to question the reasoning behind these recipes. Then, as per Jaynes:

Whenever a real scientific problem arose that was not covered by the published recipes, the scientist was expected to consult a professional statistician for advice on how to analyze his data, and often on how to gather them as well. There developed a statistician–client relationship rather like the doctor–patient one, and for the same reason. If there are simple unifying principles (as there are today in the theory we are expounding), then it is easy to learn them and apply them to whatever problem one has; each scientist can become his own statistician. But in the absence of unifying principles, the collection of all the empirical, logically unrelated procedures that a data analyst might need, like the collection of all the logically unrelated medicines and treatments that a sick patient might need, was too large for anyone but a dedicated professional to learn.

Jaynes's statements that "deep change in the sociology of science – the relationship between scientist and statistician – is now underway" and that "each scientist involved in data analysis can be his own statistician" seem premature. My impression is that basic courses in applied statistics are routinely taught without even attempting to impart much conceptual understanding, and for many scientists doing your own statistics is still dangerously close to rolling your own crypto. Be that as it may, hardly anyone can be against getting scientists to understand the statistics they are practicing. According to Jaynes, one of the earliest attempts to do this is the 1939 "Theory of Probability" by (future Sir) Harold Jefferys.

This book is perhaps the most direct prior influence on Jaynes and on PT:TLoS - which is, after all, "dedicated to the memory of Sir Harold Jeffreys, who saw the truth and preserved it". In Jaynes's telling, Jeffreys "was buried under an avalanche of criticism which simply ignored his mathematical demonstrations and substantive results and attacked his ideology". Jaynes writes:

We need to recognize that a large part of their differences arose from the fact that Fisher and Jeffreys were occupied with very different problems. Fisher studied biological problems, where one had no prior information and no guiding theory (this was long before the days of the DNA helix), and the data taking was very much like drawing from Bernoulli's urn. Jeffreys studied problems of geophysics, where one had a great deal of cogent prior information and a highly developed guiding theory

(all of Newtonian mechanics giving the theory of elasticity and seismic wave propagation, plus the principles of physical chemistry and thermodynamics), and the data taking procedure had no resemblance to drawing from an urn. Fisher, in his cookbook defines statistics as the study of populations; Jeffreys devotes virtually all of his analysis to problems of inference where there is no population.

But just in case you had any doubt whose side he is on, Jaynes then adds:

What Fisher was never able to see is that, from Jeffreys' viewpoint, Fisher's biological problems were trivial, both mathematically and conceptually.

Them's fightin words!

Incidentally, Jaynes does not deny Fisher's mathematical cleverness, and credits him with having a "deep intuitive multidimensional space intuition", which allowed him to calculate many sampling distributions for the first time. But Jaynes points out that "just before starting to produce those results, Fisher spent a year (1912–1913) as assistant to the theoretical physicist Sir James Jeans, who was then preparing the second edition of his book on kinetic theory and worked daily on calculations with high-dimensional multivariate Gaussian distributions". Yes, even these stem from a physicist whose last name starts with J!

A secondary antagonist is [William Feller](#), the author of "the most successful treatise on probability ever written". He is also accused by Jaynes of being too clever - and thus being able to get away with not doing things systematically. According to Jaynes, Feller's readers "get the impression that: (1) probability theory has no systematic methods; it is a collection of isolated, unrelated clever tricks, each of which works on one problem but not on the next one; (2) Feller was possessed of superhuman cleverness; (3) only a person with such cleverness can hope to find new useful results in probability theory". The unstated implication here is that we should doubt all three. As an illustration of "clever tricks" Jaynes chooses the following problem:

Peter and Paul toss a coin alternately starting with Peter, and the one who first tosses 'heads' wins. What are the probabilities p , p' for Peter or Paul to win? The direct, systematic computation would sum $(1/2)^n$ over the odd and even integers:

$$p = \sum (1/2)^{(2n+1)} = 2/3 \quad [n=0, \dots, \infty], \quad p' = \sum (1/2)^{(2n)} = 1/3 \quad [n=1, \dots, \infty].$$

The clever trick notes instead that Paul will find himself in Peter's shoes if Peter fails to win on the first toss: *ergo*, $p' = p/2$, so $p = 2/3$, $p' = 1/3$.

The "*ergo*, $p' = p/2$ " is saying that Paul will win precisely when [Peter does not win immediately] and [Paul wins, given that Peter does not win immediately]. The probability of the first clause is $1/2$, and that of the second is p (since, after Peter tosses a tail, Paul's situation is the same as that of Peter at the start of the game); *ergo*, $p' = (1/2)*p = p/2$. One can also solve this problem by saying that for Peter to win, he needs to either do so immediately, or later - after the first two tosses come up tails and the game effectively begins anew. In math, this says that $p = (1/2) + (1/2 * 1/2) * p$. Here, $1/2$ is the probability of Peter's immediate win, $(1/2 * 1/2)$ is the probability of two first tosses being tails, and p is the probability of Peter winning once the game starts anew, after the first two tails are tossed.

Of course, Jaynes himself can do things that are clever. His dexterity with generating functions, transform methods, and asymptotic expansions, among other things, can appear magical to those not trained as applied mathematicians or physicists. But the irony here is that this "Peter and Paul problem" is exactly the wrong example to use for complaining about "isolated clever tricks and gamesmanship". In fact, thinking about a game as a system moving between states, and analyzing how likely it is to reach certain "goal states" is basically using [Markov chain](#) theory! It is one of the most common methods for solving probability problems, well connected to other key things in probability and statistics.

This "Peter and Paul" mishap serves as an illustration of a deeper point: once well understood, many clever tricks become powerful methods, much more powerful indeed than straightforward but uninspiring computations. I agree with Jaynes in calling for "general mathematical techniques which will work not only on our present problem, but on hundreds of others". Trouble is, a 'general technique' may solve a given problem, but not explain what is going on (mathematician Paul Zeitz calls this "How vs. Why"). At the same time, a clever trick may lead to a better general theory, closer to answering the 'why' question. I am arguing not for "gamesmanship", but for bringing the game to the next level.

There are many other things Jaynes has to say about "orthodox" statistics and statisticians. One such volley is aimed at Jerzy Neyman, who had an argument with Jeffreys. In this argument, Jaynes says, "Jeffreys is clearly right". This affirmation is the only reason I see for bringing up this episode in the book, since the actual nature of the dispute is not given explicitly. What is my reason for bringing this up in the review? Well, having read the relevant parts of the [original sources](#), I can report that Jeffreys was clearly wrong. I encourage you to discuss whether I am wrong that Jaynes is wrong that Neyman is wrong in the comments.

In Persi Diaconis's [review](#) I mentioned earlier, he calls PT:TLoS "wonderfully out of date", saying that "the wonderful part is that Jaynes discusses and points to dozens of papers from the

1950s through the 1980s that have slipped off the map." A noticeable fraction of this pointing is, in fact, pointing fingers at people doing things wrong. The abundance of these sidetracks forces the reader to either mostly ignore them or to follow up on them. Both strategies are admissible, and I have found the second one quite rewarding, but it makes reading PT:TLoS seem like walking through a garden of forking paths.

Paradox lost.

Jaynes's opinions are of course not limited to statistics. He has things to say about set theory, measure theory, the infinite, Kolmogorov's axiomatization of probability, generalized functions, Godel's incompleteness, and so on. Jaynes says that "we shall find ourselves defending Kolmogorov against his critics on many technical points". I was glad to see this, but not because I think Kolmogorov needs defending. Rather, it signaled to me that Jaynes's math will be mostly right. Despite this, after I read Appendix B, it became clear to me that on the subject of modern mathematics Jaynes and I don't really see eye to eye. This appendix contains most of Jaynes's attack against modern mathematical formalism, but remarks of similar nature can be found in multiple places in PT:TLoS. I view finding the right language and level of generality, and using mathematical rigor to remold uninformed intuition as some of the essential goals that underpin modern mathematical developments. Unfortunately, they seem to be either ignored altogether, or viewed as unimportant hindrances by Jaynes. He also insists that using modern techniques can produce nonsense. In fact, one can produce nonsense even without any modern math, simply by not being careful. Perhaps you have seen how one can use the sum $S = 1 - 1 + 1 - 1 + \dots$ to show that $0 = 1$: namely, $S = (1 - 1) + (1 - 1) + \dots = 0$, but $S = 1 + (-1 + 1) + (-1 + 1) + \dots = 1$. You should not interpret this as saying that using arithmetic can produce nonsense. Rather, you should learn that unjustified manipulations like that are dangerous - as indeed they are. That's precisely why mathematicians have thought long and hard about how one can work with such infinite series without running into problems. They developed multiple sophisticated and precise theories about this, some of which are now taught in the 'sequences and series' part of courses on rigorous mathematical analysis. Ignoring what these theories say is what leads to apparent paradoxes.

The situation with paradoxes of probability theory is quite similar. This might explain why chapter 15, "Paradoxes of probability theory", was not all I hoped for. A true mathematical paradox would be a pair of contradictory conclusions derived from standard axioms. So far, no such thing has been found. Thus, any presently known "paradox" must be of some other nature. Roughly speaking, there are three common types. There are true statements that subvert naive intuition (a la Banach-Tarski paradox), there are faulty demonstrations (like Achilles and the

tortoise), and there are arguments that reveal a deficiency of terminology or definitions (such as Russell's paradox). Alas, many of the "paradoxes" in PT:TLoS are not even paradoxes in this weaker sense. For example, consider the following: we take one of the Kolmogorov's axioms called 'countable additivity' and weaken it to something called 'finite additivity'. We study the resulting alternative "probability theory" and find that it sometimes produces pathological-looking results. Is this a paradox? I'd say hardly so. Yet, this is exactly what the "non-conglomerability paradox" from chapter 15 boils down to after all is said and done.

Another example, the "Borel-Kolmogorov paradox", is mostly of terminological type - it poses the question of how to make sense of conditioning on an event of probability zero. Jaynes considers the case of conditioning a joint density. Then, he shows that a plausible-looking formula for this conditioning is easily obtained "by an intuitive ad hoc device". Jaynes shows how careless use of this formula gives rise to paradoxical results. Surprisingly, Jaynes does not point out that this paradox was resolved by Kolmogorov in the same 1933 [book](#) where he axiomatized probability theory. Using some basic measure theory Kolmogorov defines conditioning with respect to a random variable, of which conditioning a joint density is a special case.

Finally, the "marginalization paradox" touches on an important issue of using improper priors in Bayesian inference. Roughly speaking, 'improper' means 'not summing to 1'. It's as if you added up probabilities of every possible outcome, and found that the sum, i.e. the probability that something will happen, is not 1, but instead is infinite. Clearly, this is not supposed to happen. Using such priors voids all the warranties on your calculational procedures, and can lead to contradictory answers. So why not just agree to never use them? Because they are often easier to compute with, and, more importantly, they arise often from both maximum entropy and the group invariance we have talked about. So Bayesians would like to know how to work with them. The marginalization paradox is an example of a problem in which two different approaches to using and improper prior do in fact produce different answers. Jaynes's solution is to approximate improper priors by proper ones, perform the necessary calculations with those, and then take the limit. The trouble is that the calculations are sufficiently involved, and the limits are sufficiently tricky that following everything to a satisfactory conclusion is challenging. Even a very persistent reader may end up [doubting](#) Jaynes's conclusion as to which of the two procedures should ultimately be used.

Exegi monumentum.

What are we to make of all this, [as the saying goes](#)?

PT:TLoS is, to put it mildly, a very special book. It is neither a textbook, nor a reference text, nor a philosophical treatise, nor a history book - and it is a bit of all of those. It is singularly shaped by the person of E. T. Jaynes: by his "two thirds theoretical physicist and one-third electrical engineer" background, by his interest in radars and statistical mechanics, by his unconventional thinking, by his polemic style in disputes with statisticians of his age, and by his untimely death.

The book's chapters written earlier and polished for longer are some of the strongest, while those added late are often more susceptible to criticism or are incomplete. Yet, despite its flaws, the book had a tremendous influence in reframing statistical practice through the lens of inference and promoting Bayesian methods as a coherent framework for it. Bayesian methods have become ubiquitous in machine learning, and maximum entropy distributions (often under the name of [exponential families](#)) play a prominent role. Perhaps because of this influence PT:TLoS is often recommended as a resource on probability and information theory for those with "absolutely no prior experience with these subjects" or even "to the general reader". I find this akin to recommending "Ulysses" as a practice book for beginner English learners. PT:TLoS would fare much better as "A Companion to Probability: [A Second First and A First Second Course](#) in Probability", suitable for a dedicated reader possessing a solid grasp of the basics and wishing to gain a deeper conceptual understanding. An advanced student or a practicing professional in any field related to probability, statistic, or machine learning would benefit from reading it. As it stands, however, PT:TLoS is one of those complex classics that many wish to have read, but not many have actually managed to read.

Perhaps this review may at times seem critical and not sufficiently expounding on all of Jaynes's contributions. This may be because Jaynes has, by now, won many of his battles, and his mode of thinking has become part of the intellectual background of our age. It is, after all, difficult to appreciate an insight once it becomes the usual mode of thinking, the [proverbial water](#). It may also be because the book itself is incomplete, and sometimes frustrating. In the very first paragraph of the editor's preface, Larry Butterhurst explains:

I could have written [the] latter chapters and filled in the missing pieces, but if I did so, the work would no longer be Jaynes'; rather, it would be a Jaynes–Butterhurst hybrid with no way to tell which material came from which author. In the end, I decided the missing chapters would have to stay missing – the work would remain Jaynes'.

This is a decision which one [Amazon review](#) calls "a bad mistake". This is certainly how I felt when I was reading the book; now I am less sure. Once you have struggled through it, the motivation to make the struggle less onerous diminishes, and you begin to think that "keeping

the work Jaynes'" may actually be a valid consideration, and not just a lazy cop-out you thought it to be whilst in the thick of it all.

And yet I, too, find myself mourning for what this book could have been. Sometimes when faced with a choice (ketchup or mayo? vanilla or chocolate?) I simply choose both. We already have the Jaynes's version of PT:TLoS. Can we not get the "completed version" as well? Could we not write the missing chapters, explain the cryptic references, solve the unsolved exercises and release the result to the world? Someone who is better than I am at organizing things, and someone who knows more than I do about copyright and publishing would need to think about it. On one hand, we are in the 21st century, with the power of the internet, crowdsourcing and social campaigns. On the other hand, it is my understanding that it will almost be the 22nd century before the copyright for PT:TLoS expires.

Until then, we read the version we have. The version that embodies Jaynes's message: "progress in science goes forward on the shoulders of doubters, not believers". The version that urges us to think for ourselves rather than to defer to the "orthodoxy" (whatever it may be called in our time), to see the truth and preserve it.

Reasons and Persons by Derek Parfit

"Like my cat, I often simply do what I want to do."

Derek Parfit

Unfortunately, people are not cats and we expect them to have reasons for performing certain actions, be those reasons moral, rational, or otherwise. The reasons we have for acting in a specific way is one of the major questions that Derek Parfit attempts to answer in *Reasons and Persons*, with the other being what is the nature of personhood. The book is divided into 4 parts, with parts 1 and 2 focused on reasons for actions, part 3 on the nature of personhood, and part 4 on the repugnant conclusion. I find Parfit's arguments largely convincing, and the book is quite readable for a 500 page philosophy text. Others may be more shocked by Parfit's conclusions and find them difficult to accept. Parfit believes that persons are fictional, or at least that personhood does not matter in any important sense. Parfit is concerned with personhood in the sense of "am I the same person I was as a child?" not in the sense of "are dolphins persons?" He also believes that in some cases determining whether someone is the same person as someone else is impossible and/or meaningless. The organization of the book is somewhat scattered; he will go off on tangents that will slowly work their way back to the main point of the chapter, but it usually easy to follow along. He also has nice summaries at the end of each section that are well organized and help re-focus the previous discussions on his central points. His jargon and theories are easy to understand, though it can be difficult to keep track of which is which when there are theories S, P, Q, and others floating around. I have read the book multiple times, which definitely helps in understanding and picking up smaller things you might have missed the first time. I would strongly recommend the book to anyone interested in ethics, personhood, or rationality. I would still recommend it, though less strongly, to anyone else.

Part 1 is probably the least interesting part of the book and the least organized, though not without its merits. Parfit frames the discussion in part 1 around a theory of rational self-interest that he calls S. S states that each person's ultimate aim is for their life to go as well as possible. In other words, what is rational to do or what you have the best reasons to do is whatever would make your life go well. Although there is a lot of ambiguity in what "as well as possible" means, it does not affect Parfit's discussion too much. Another theory of rational self-interest Parfit discusses is P. P states that each person has most reason to do whatever will fulfill their current goals and desires. Unlike S, P is only concerned with the present.

One of the more interesting side discussions that arises from Parfit's evaluation of S is collective consequentialism, which sort of combines Kant's universalizability (you should act in such a way that you could rationally wish that everyone would act in the same way) with utilitarianism. This moral theory states that one should act in such a

way that if everyone acted in the same way, utility would be maximized. For example, under regular utilitarianism you might figure that you would maximize utility by donating 90% of your salary to people suffering from global poverty. This might make you very unhappy, but so many people in the 3rd world are helped that total utility is maximized. Under collective consequentialism you would calculate instead what would be the ideal amount for someone to donate in your position if everyone in a similar position donated the same. This number would be much lower than 90%, because in this ideal situation many more people are donating. It might be something like 50%. Although the world as it is would have less total utility by collective consequentialism, it has two major advantages over regular utilitarianism. It is more fair, in the sense that everyone has the same obligations and the fact that other people are awful and are not doing their fair share no longer means that you have to sacrifice even more to make up for that (If everyone was donating more in the first example, the ideal amount for you to donate would be less than 90%. For each person that opts out, the amount that you should donate goes up). The other advantage is that calculating utility becomes easier because you get to assume an ideal world. You no longer have to figure out how each person would act in the real world; you get to assume everyone acts the same and acts morally. Obviously, these assumptions are not going to be accurate in the real world, but at least it makes the very difficult problem of actually sitting down and calculating utilities of various actions more tractable.

Part 2 builds off part 1 and argues that S is not the correct theory of rationality. Morality would indicate that you have reasons to be concerned with other people, not just yourself, so S needs to make a case that ignoring other people in a moral sense is rational. However, S also needs to make the case that ignoring your future self is not rational, otherwise P would be true, not S. In other words, if you don't care about other people, why should you care about your future self? The answer to this question may seem obvious, namely that you will never be other people but you will be your future self, but Parfit undermines the idea that one is the same person existing throughout one's whole life in part 3. (There is a lot more discussed in parts 1 and 2, such as Parfit's namesake hitchhiker, but I found part 3 to be by far the most interesting part of the book, so I am drastically simplifying and compressing the first 2 parts).

In part 3 Parfit discusses personal identity in what is likely the meatiest and most famous part of the book. He begins with the transporter thought experiment. A machine scans your body, and makes an exact replicate (quantum physics notwithstanding) on Mars while destroying the original. Parfit considers several alternative variants as well, which result in your original body being fine or your original body being sickened and soon dying instead. Which of you (or both or neither) is "you" in each instance? If you died, is your replicate "you?" How concerned should you be if the process sickened you and you will die shortly if you have a healthy replicant elsewhere? If you survived, are you both you? Is there a theory that can explain what happens in all 3 scenarios satisfactorily?

Parfit evaluates several theories of personal identity. Physical continuity is one possibility. You are the same person over time if you keep the same body. In this case

none of the copies are you in the transporter thought experiment, because your body is still on Earth in whatever state it is in and your replicant is on Mars. The degree to which your body needs to be the same is debatable, but there must be some part that is identically or at least continuous with the other parts. Psychological continuity is another theory, one that is based on continuous memories instead of a continuous body. I remember being myself yesterday, yesterday I remember being myself the day before, therefore I am the same person today as I was 2 days ago. In this case the copy will be you, although it does run into some confusion when there are multiple people who are or were psychologically continuous with each other running around. Parfit will argue for a version of the psychological continuity theory of personhood. These are both reductionist views, in that personhood is merely used as descriptions of a set of facts. Parfit believes that being a reductionist about personhood compels you to believe that personhood is unimportant, but other philosophers would disagree. There are also non-reductionist views on personal identity, such as that having the same soul over time makes you the same person. The soul exists independently of mere facts about the body or mind. Parfit also suggests that there can be materialist non-reductionist theories about personal identity, but I don't know if there really can be.

For example, most people are reductionists about nations. Nations clearly exist, and consist of people and/or territory, but there is no deeper fact or soul of a nation beyond the people or territory. In addition, reductionists will be open to the fact that in some cases we will not and cannot know if the object in question is the same "thing" or not as it was before. Whether the Austrian Empire is the same country as the current Austria does not have a "true" answer. There may be arguments for saying that it is and arguments for saying that it is not, and it is possible to reach a consensus on the question, but there is no logical proof one way or the other. This seems all well and good in regards to nations, but in terms of personhood, being a reductionist means that questions like "Am I about to die?" and "Will I be the same person a year from now that I am today" may be answered "Maybe" "If you want to be" or "I have no idea." Parfit holds that this is the correct view towards personhood. Going back to the teleporter experiment, there is no "true" answer to the question "am I about to die?" according to Parfit. There will no longer be a body and mind like mine on Earth, and there will be one on Mars. That's it, that explains the whole story. You can call the one on Mars me or not, it's just a matter of preference.

Under normal circumstances, answering "I don't know" to the question if someone is the same person or not is absurd, but Parfit proposes another thought experiment that would make it appear to be a reasonable answer. If a mad scientist replaced your memories and personality slowly with that of another person, at what point do you "become" a new person? Answering that you are definitely yourself early in the process, definitely someone else later in the process, and you don't really know who you are in the middle of the process seems reasonable, at least compared to the alternative were changing one tiny memory or trait changes you from 100% yourself into 100% a new person. Although this experiment has never been done and probably will never be done, we do have a lot of evidence from people suffering from brain trauma

and mental illnesses that you can lose a piece or pieces of your memory or personality while keeping the rest intact (See Oliver Sack's *The Man who Mistook his Wife for a Hat* for some extreme examples). Whether or not these are the same people as they were before is a difficult question, which Parfit would argue has no right answer. If instead we found that brain trauma would either do nothing or completely rewrite everything in your mind, that would have been evidence for a non-reductionist view of personal identity.

For another example where there appears to be no answer to if you are the same person or not, Parfit proposes that your brain could be split and transplanted into two separate bodies that belong to your other identical triplets. Parfit assumes that the human brain in this thought experiment has 2 independent and nearly identical hemispheres. I do not know if current neuroscience would support these facts. His argument maintains at least some force as long as the scenario is physically possible, even if it could never be done with human brains. Although you would lose your own body, both new people would be very similar to your original self. Any possible answer to the question which, if any, of the 2 new people is the same person as the original doesn't work. That neither of them is the same person doesn't make sense because people lose significant parts of their brain to strokes and the like and remain the same person as they were before. Your body is nearly identical to your original body in this scenario, and in real life people's bodies change drastically over time. It would be impossible to select only one of the two as you, because they are identical after the surgery (or at least they are equally and mostly identical to the original). There could be no possible basis for choosing one over the other. Calling both of them the same persons seems to undermine some basic facts about personhood, like that persons are unique and do not have multiple independent bodies (I would be curious to see a theory of personhood that allows for multiple independent bodies). We have all the information, but still no idea if the same person still exists over time. From the point of view of the person whose brain was transplanted, both new "persons" would still have most of memories, personalities, skills, etc and definitely feel like they are still the same person as before.

"There will be two future people, each of whom will have the body of one of my brothers, and will be fully psychologically continuous with me, because each has half my brain. Knowing this, we know everything. I may ask, "But shall I be one of these two people, or the other, or neither?" But I should regard this as an empty question. Here is a similar question. In 1881 the French Socialist Party split. What happened? Did the French Socialist Party cease to exist, or did it continue to exist as one or the other of the two new Parties? ... Even if we have no answer to this question, we could know just what happened."

Parfit argues that division in this manner, as well as teleporting, is just as good as ordinary survival. Since questions of personhood are empty, the fact that the "people" in the scenarios have the same memories and personalities as before, and nothing feels

that different internally, they are fine. Personhood does not matter; maintaining thoughts, goals, feelings, etc is what is important. If we consider these scenarios “from the inside” nothing seems wrong. Nothing important is missing. If you were to feel different, have confusing memories or a sudden change in personality, that would be a sign that something significant happened, but that is not the case in the relevant though experiments. I simply step into the teleporter or lay down on the operating table, briefly fall unconscious, and wake up on Mars or on a different operating table. It would feel no different than waking up from anesthesia after surgery. You wouldn’t notice anything wrong or missing. To say that you “die” when going through the teleporting is to say you die when you go under anesthesia. You could be destroyed and replicated constantly, and it would both make no difference and would be unnoticeable.

“I want the person on Mars to be me in a specially intimate way in which no future person will ever be me. My continued existence never involves this deep further fact. What I fear will be missing is always missing... Judged from my standpoint of my earlier belief, this not because teletransportation is about as good as ordinary survival. It is because ordinary survival is about as bad as, or little better than, being teletransportation. Ordinary survival is about as bad as being destroyed and replicated.”

This is the most important idea in the entire book, and one of oddest and most profound ideas I have ever encountered. Ironically it is almost identical to the Buddhist idea of no-self, which Parfit also acknowledges, although he arrived at in a very different fashion. Parfit even describes this realization in a way that if you squint enough looks like Enlightenment.

“I find it liberating and consoling. (Before) I seemed imprisoned in myself. My life seemed like a glass tunnel, through which I was moving faster each year, and at the end of which was darkness. When I changed my view, the walls of the glass tunnel disappeared. I now live in the open air. There is still a difference between my life and the lives of other people. But the difference is less. Other people are closer. I am less concerned about the rest of my own life, and more concerned about the lives of others.”

This seems like it should have profound consequences, which Parfit acknowledges though downplays. Does it make sense to do things like save for retirement, punish people and hold them accountable for acts they committed in the distant past, or even to keep promises or commitments made in the past if we can’t say we are the same person over time? Is worth it to start smoking now to look cool even if “I” get lung cancer 30 years from now? Should we even be afraid of death, if we are being “destroyed and replicated” all the time? Is there any reason to favor ourselves or our future selves over other people? Parfit does hold that memories, personalities, and goals are important and provide reasons for considering the future and not wanting to

die, although we should be less concerned about these things than most people are now. I worry though that this is not enough; many people have completely different personalities and goals at 50 than at 20, and have forgotten a lot of memories and made a lot of new ones. How strong is this relationship between your 20 year old self and 50 year old self? And can it justify things like not dating that hot girl now because she would be a terrible mother and wife in 20 years? And if you barely resemble your 20 year old self at 50, should your 20 year old self regard this change as being as bad as dying? I apologize for the question overload, but these are things that I think we must think about if we take Parfit seriously, and I don't really know how to answer them myself.

Parfit holds that it could be rational to give our future selves no special concern, although giving our future selves special concern is permissible as well. The best Parfit can do is give a moral argument for prudence. No one will be in a better position or have a greater ability to help your futures self than you, so you have a moral duty to do so, just like you have a lesser moral duty to help other people. You could also make the argument that you have a special duty towards your future self like you have a special duty towards your spouse or children. This is a lot weaker than traditional arguments for prudence, but at least it is something.

This argument has other interesting effects; we often don't morally criticize people for actions that hurt only themselves, but by this view we should treat these kinds of actions in the same way we would if they were hurting someone else. By the same argument, paternalism and restricting others' freedom for their own good would be easier to justify, and hurting someone for their own future benefit would become harder to justify. It also seems like an argument for effective altruism; how can you justify spending millions on your future self's retirement instead of on people living in extreme poverty? How can the benefit of one person in the future (who isn't you and at best you have some vague responsibility for) outweigh benefitting hundreds of much needier people now?

In regards to punishment Parfit similarly believes that punishment is not required but is still permissible under his view of personhood. He does not make many definitive claims, except that as someone changes over time away from the person they were when they committed the crime, they deserve less punishment, although perhaps not none. His example is that an elderly Nobel Peace Prize winner should not be punished if it comes to light that he started a drunken brawl in his 20s, even if it would have been correct to punish him at the time. This probably doesn't fall too far out of the average person's moral intuitions, but I can imagine more bizarre examples that would. And since Parfit is such a big fan of crazy sci-fi thought experiments, I'm sure he would not have minded if I do one myself. Imagine a killer is on the run and is about to be captured and severely punished. However, he has the ability to instantly change his personality and forget all details of his crime. He and no one else knows if the changes will be permanent. If he does changes himself, the police can't punish him, since he is no longer connected in any important way to the criminal (according to Parfit). Maybe they could detain him until his original personality returns, but the longer this would take

the less plausible it would become. We don't preemptively detain people except in extraordinary circumstances, and imprisoning an innocent man because at some point he might become someone dangerous or someone who should be punished seems really weak. Theoretically anyone might become someone dangerous at some point in the future. It kind of becomes a get-out-of-jail free card if you are willing and able to drastically change yourself after committing a crime. (This thought experiment was inspired by *Memento* and *Death Note*). Parfit might be fine with this, as he considers drastically changing yourself in such a way to be just as bad as dying, and suicide already functions as a get-out-of-jail- free card of sorts. I don't know how he would account for the chance that you would revert to your original self though. Parfit readily admits that connectiveness over time is something that you can easily influence, so you could choose to give yourself greater or less unity over time by how much you change yourself, although he does not consider anything as drastic as my thought experiment.

Part 4 deals with several things, but the primary one is the Repugnant Conclusion. The Repugnant Conclusion states that by utilitarianism, it would be better for a vastly greater number of persons to have lives just barely worth living by whatever measure of utility you chose than for far fewer people to have much better lives. In mathematical terms, a situation with 1 million lives each having a net utility of 1 is better than a situation where 1,000 people each have lives of net utility 100. $1,000,000 \times 1 > 1,000 \times 100$. This conclusion can also apply over the course of a single life. Is it better to have a very long but mediocre life or a short amazing one? 100 years of utility 1 is greater than 30 years of utility 3. Parfit finds this situation counter-intuitive and disturbing, hence the name. On a smaller scale, questions of whether or not to have children (or have an abortion) follow similar logic. How many children is the "correct" number to have? As many as you can while maintaining their (and your) quality of life at a level that is worth living? Or the amount that will maximize the existing children and your quality of life, which is likely a much smaller number?

Parfit attempts a number of mathematical tricks to avoid the Repugnant Conclusion, including that quantity of people shouldn't matter above a certain point, that suffering shouldn't matter beyond a certain point, that happiness shouldn't matter beyond a certain point, and that some very good lives might be infinitely more valuable than any number of lives barely worth living. He finds all of them implausible, since one could construct a similar thought experiment to the Repugnant Conclusion for each of them which is even more absurd and/or repugnant. The more obvious solution would be to reject utilitarianism, which Parfit does not consider. I am struggling to imagine how other moral theories would even tackle the question of which world is better. Which world does not use people as mere means? Which world would the man of virtue chose to bring about if he had the choice? (Maybe the smaller scale question of how many children to have could be answered by one of these theories, and then extrapolated). While worrying, I don't think the Repugnant Conclusion is a knockout argument against utilitarianism, and I am not as repulsed by it as Parfit. Until we have a better handle on what exactly a life that is worth living, but only barely, would even look like, I don't think it's worth spending too much time on.

I would recommend this book to everyone reading this review. My complaints are ultimately minor compared to how enjoyable and enlightening the book is. The most serious of these complaints is that Parfit does not seem to show proper concern for all the bizarre implications that his theory of personhood would have. I would have liked to see more definitive ideas in this section, or an acknowledgement that the lack of any is a major problem. A more minor disappointment was that the question about whether your blackout drunk self is the same person as your sober self was never raised (I have been puzzling over this since I saw a BuzzFeed post where a person left a glass of vodka out while blackout drunk with a note saying it was water to prank his sober self the next morning. I cannot decide whether this works as a prank or not). I have omitted a lot of tangential content which was still interesting, and Parfit did a much better job explaining these concepts than I did. I am somewhat biased though, since I have a background in philosophy and Buddhism, I found Parfit's arguments surprisingly easy to understand and accept, I imagine others might find him far less convincing or far more disturbing than I did. The idea of personhood being of extreme importance is probably very useful evolutionarily; it encourages activities like planning for the future and valuing your own life over others. It should not be surprising that it is difficult to discard, even if one accepts the idea that we should. I will end with the same quote from the Buddha that Parfit ends with:

“O Brethren, actions do exist, and also their consequences, but the person that acts does not. There is no one to cast away this set of elements and no one to assume a new set of them. There exists no Individual, it is only a conventional name given to a set of elements.”

Reinforcement Learning: An Introduction by Richard S. Sutton and Andrew G. Barto

On the surface, this is a technical textbook for those wishing to learn about reinforcement learning (RL), the subfield of machine learning algorithms powering recent high-profile successes of AIs trouncing human champions like AlphaGo in Go and Watson in Jeopardy.

However, in this review I want to talk about something quite unexpected: *Reinforcement Learning* profoundly changed my understanding of the science of happiness, biological evolution, human intelligence, and also gave me unique tactics for rapid skill acquisition in my personal life.

For brevity, I will focus on highlighting only the most important (and underrated!) topics from this book that you won't find elsewhere. If you want a comprehensive summary of the book, you can instead read my detailed notes on each chapter here: <https://billmei.net/books/reinforcement-learning/>

What is reinforcement learning

Reinforcement learning is a study of what are the best actions (*policy*) to take in a given environment (*states*) to achieve the maximum long-term *reward*.

I've found RL to be the most elegant, first-principles formalization of what it means to "win". It's an algorithm that explicitly gives you the optimal actions to take to achieve the goal you define, with zero outside knowledge other than the input environment. Unlike other machine learning algorithms, RL does not require you to specify subgoals (capture the most chess pieces), only the ultimate goal (win at chess).

The SSC and LessWrong communities started using Bayes' theorem as the principal guide towards finding the truth. Bayes' theorem is the first-principles formalism of how to evaluate evidence and update your beliefs so that your beliefs will match reality as closely as possible. But ultimately "[Rationality is winning](#)"; the purpose of "Rationality" is not just to have a good epistemology, but to successfully achieve goals in the real world.

Just as Bayes' theorem is the mathematical foundation for finding the truth, reinforcement learning is the computational foundation for winning.

How reinforcement learning works

To oversimplify, let's focus on temporal-difference (TD) learning. TD learning uses a *value function* that estimates how good it is to take a certain action, or how good it is to be in a *state*: a certain configuration of the environment. For example, in chess it's generally better (but not

always) to have more pieces than your opponent. It's also generally better (but not always) to take actions that capture pieces than to take actions that result in your pieces being captured.

If the value function predicts that a state is better than it actually is, causing a move that results in losing the game, then we perform one iteration of a *TD update* which revises the value function downwards on this particular state (and other similar states). Vice-versa if it undervalued a state that later resulted in a win. We call this misestimation a *TD error*.

The RL agent will then mostly try to play only the actions that lead to board states with the highest estimated value, although it will sometimes play suboptimal actions to avoid getting stuck in local optima using a strategy that I will describe later.

Over many games and many iterations, the value function's estimate asymptotically converges to the true value of the state (played by some hypothetically perfect player). Then, assuming these estimates are correct, the RL agent can win by simply playing actions that lead to the highest value states. Even if the estimated value does not perfectly match the true value, in most practical cases the estimates become good enough that the agent can still play with superhuman skill.

Mistakes drive learning

What's interesting about the TD algorithm is that it learns only from its mistakes. If the value function estimates that I will win a game, and then I go on to actually win, then I get a positive reward but the TD error is 0, hence *no TD update is performed*, and no learning occurs. Therefore, an RL agent that learns quickly won't always just choose actions that lead to the highest state values, but instead identify states that haven't been played frequently and try to play in such a way to get to those states to learn about them. While this may cause the agent's winrate to decrease in the short term, it's a good way to improve out of local optima as focusing only on a small number of states leaves the agent ignorant about the universe of what's possible.

This echos the research on how you can acquire skills using "deliberate practice". The popular "10,000" hours rule oversimplifies the idea of "just practice a lot and you will be good", as the research of *deliberate* practice shows that if you're just doing mindless repetitions of an activity (or worse, winning a lot or having fun at it!), you aren't actually learning, as it requires struggling through mistakes and constantly challenging yourself at a level just beyond your reach that results in actual learning.

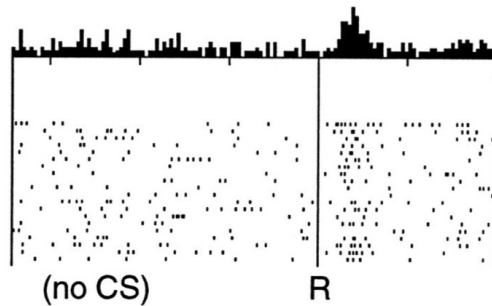
While you may already be familiar with the research on deliberate practice, RL provides a *mathematical* justification for this concept which further solidified it for me.

Dopamine and Happiness

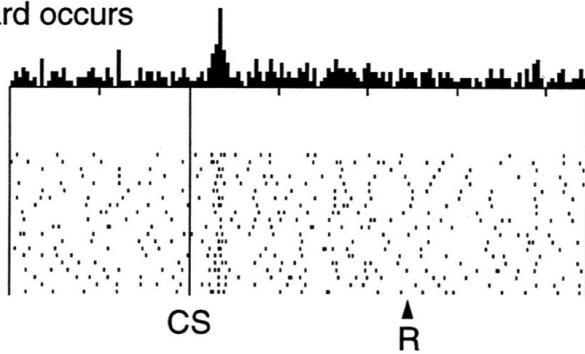
Dopamine is a neurochemical associated with pleasure. While it may seem like dopamine is analogous to the *reward* in an RL algorithm, it is not the reward. Instead, dopamine is the *TD error*.

Echoing Pavlov's dogs, the book describes a study where researchers (Schultz, Apicella, and Ljungberg, 1993) trained monkeys to associate a light turning on with a reward of a drop of apple juice. They hooked up brain scanners to the monkeys to monitor their dopamine receptor activity. Here's how those neurons behaved:

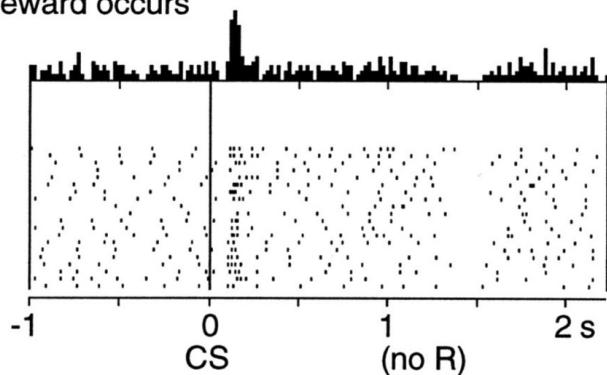
No prediction
Reward occurs



Reward predicted
Reward occurs



Reward predicted
No reward occurs



The x-axis is time elapsed and the y-axis is strength of dopamine response. CS: Light turning on
R: Drop of apple juice dispensed

From the diagram, the monkeys get a dopamine spike when the apple juice is surprisingly dispensed without any forewarning from the light. If the light turns on, the monkeys get a dopamine spike right away in anticipation of the upcoming reward, but then *no dopamine is produced once the apple juice is actually dispensed*. In a third experiment, the light turned on which produced a dopamine spike as expected, but then later no juice was dispensed, which caused the monkey's dopamine levels to drop below baseline as their expectations were disappointed.

From this it's clear that dopamine's purpose is an error signal, not a reward signal. We don't get any dopamine when our rewards are exactly the same as our expectations, only when the rewards exceed expectations. Conversely, dopamine levels drop when our expectations are high and the rewards are disappointing. Likewise, once an RL agent receives a reward its TD error is positive when the value function undervalued its actions, and negative when it overvalued its actions.

It's a trope that "happiness equals reality minus expectations", and while dopamine is not the only neurochemical that contributes to happiness, the implication of this study is the more skilled you get at accurately predicting reality, the less pleasure (and less disappointment) you get from your efforts. Perfect prediction is perfect equanimity.

Another implication is in the psychology of addiction. This phenomenon underlies the behaviour of "chasing the high"; every time you receive a reward, your expectations revise upwards to match reality, so next time the reward needs to be greater to generate the same dopamine response, as receiving the same reward causes no "TD error".

These conclusions may be unsurprising to you if you are a psychiatrist, but what I found extraordinary is the research and science around dopamine was discovered many years after when the TD algorithm was developed. The RL algorithm designers were totally ignorant about this property of dopamine, yet independently came up with an algorithm for shaping the behaviour of computer agents that looks remarkably similar to how we incentivize behaviour in animals.

Is there a shared computational universe that drives the behaviour of both biological beings and computer algorithms? It certainly looks convincing.

Simulated Experience

In chess you have full knowledge of the environment, but what about tasks where you don't? Imagine a robot trying to navigate a maze. If this were an animal study we may reward its success with food, but AIs are satisfied to receive rewards as floating point numbers, with greater rewards for faster navigation. At every intersection, the RL agent must decide which corridor to follow to exit the maze.

How can the RL agent know what “state” it’s in when it doesn’t even know what the maze looks like? Any given intersection may look exactly like several others. Worse, what happens if the maze is constantly changing while the agent is in the middle of running it?

In the absence of external sensory information about the real-world maze, the RL algorithm constructs simulated mazes that it runs virtually. Driving a robot through a real maze is slow, but the computer can run thousands of simulated mazes per second.

In the beginning, when the agent knows nothing about the real maze, its simulated mazes will be poor imitations of the real thing, as it lacks data on what mazes are supposed to look like. So the experience it gains from these simulated mazes is worthless.

But slowly, as the robot drives through the real maze and collects increasing ground-truth data about the maze, it can improve its simulated mazes until they become reasonably accurate approximations of the real thing. Now when the algorithm solves a thousand simulated mazes, it gains the experience of having solved a thousand physical mazes, even though in reality it may have only solved a handful.

This is how Monte Carlo Tree Search works (although I oversimplified here for brevity), and it was the key to AlphaGo’s victory over the top human Go players.

In the book *Sapiens*, Yuval Noah Harari argues that what separates humans from other primates is our ability to imagine events that don’t exist. Our ability to learn from these fictional events is what endows us with intelligence.

If you can daydream about a non-existent lion, then you don’t have to wait to actually encounter a lion to figure out what to do, when it may be too late.

At the risk of inappropriately anthropomorphizing our RL agent, I was convinced that this ability to simulate experience is one of the key building blocks of intelligence, here applied to machines instead of humans.

Simulated Rewards

The ability to simulate experience necessarily also means being able to imagine rewards (or punishments) at the end of those simulated experiences.

Any basic optimization algorithm suffers from the problem of being stuck in local optima. Humans can think globally and creatively because we can delay gratification; we generally don’t always take the greedy step towards rewards, as anyone who does so is seen as more simple-minded or lacking willpower. Conversely, we respect people who have superior self-control, as working towards long-term goals generally leads to more success, health, popularity, etc., thus we perceive these people to be more intelligent.

I want to argue here that our ability to delay gratification is not the result of willpower, but actually a hack. We don't really delay gratification, instead we substitute a real reward for an imagined one.

In the famous [Stanford marshmallow experiment](#), children who were able to give up a marshmallow to wait 15 minutes in an empty room received 2 marshmallows afterwards. Compared to the kids who didn't wait, the kids who waited later had improved SAT scores, educational attainment, and other measures of life outcome.

If you watch some videos of this experiment, what's remarkable is you will notice the most successful kids aren't the ones who have iron willpower, but instead those who were able to distract themselves by singing songs, playing with their hands, etc.

Thus, the key to long-term planning is not the ability to push back a reward, but instead the ability to be satisfied with an imagined fiction tell yourself of an even greater reward you can receive if you wait.

While I use the terms "fiction", "simulated", and "imagined", it's important to note that this "synthetic happiness" is not fake. Biologically, psychologically, and computationally, it is in every way as real as "real" happiness. Dan Gilbert, the happiness researcher, presents the data behind this in a [TED talk](#):

We smirk, because we believe that synthetic happiness is not of the same quality as what we might call "natural happiness". [...] I want to suggest to you that synthetic happiness is every bit as real and enduring as the kind of happiness you stumble upon when you get exactly what you were aiming for.

From our RL algorithm's perspective, its simulated rewards are as real to its value function as the real rewards. The only difference is the real rewards have the ability to change the way the simulation itself is constructed, whenever expectations are violated. The way the math works, the agent's ability to long-term plan results not from its delaying immediate rewards, but substituting real short-term rewards for simulated long-term rewards that have a larger floating point value.

Simpler animals require Pavlovian training and direct rewards/punishments to influence their behaviour, but you can get a human being to toil away in the fields with only a story about an imagined afterlife.

Importance Sampling

After three moves in chess, there are 100+ million possible board positions and 10^{120} possible continuations (more than the number of atoms in the universe), only a tiny sliver of which result in a win. To have any hope of getting a good value function with limited computing power, your algorithm must focus on analyzing the most promising (or most common) moves and avoid spending clock cycles on positions that are clearly bad and will never occur in a real game.

But your opinion on what is “good” may differ from someone else’s opinion. An RL agent’s experience in a game is highly path-dependent; if it happens to get lucky with a certain sequence of actions it may overvalue these actions and hence choose them more often than a counterfactual average agent. Thus, how much credit should you give to your own experience vs. others’ experiences? [Morgan Housel](#) says “Your personal experiences make up maybe 0.00000001% of what’s happened in the world but maybe 80% of how you think the world works.”

The *importance sampling ratio* is a modifier used by the RL agent to upregulate or downregulate its value function to reduce the variance from luck, without changing the expected value. It’s calculated using a set of weight parameters that is adjusted based on a *behaviour policy*, which you can think of as the RL agent’s simulation of what an average agent would do in its stead.

Just as Bayes’ theorem gives you the math for *exactly* how much you should increase or decrease your confidence in a belief in response to new evidence, importance sampling gives you the math for *exactly* how much credit you should give to your own experiences versus others’ experiences so you can correct for your narrow slice of reality without throwing up your hands and always deferring to others.

I believe importance sampling is the appropriate response to avoid overcorrecting and, [as Eliezer describes](#), to avoid becoming “too modest” to the point where you stop trusting your own intuitions even when they are right.

Conclusion

It may be offbeat to do a book review of a dense, 500-page math textbook, but I found the ideas presented in *Reinforcement Learning* revolutionary to clarifying my understanding of how intelligence works, in humans or otherwise.

In this review I’ve omitted the math formulas because they are quite involved and any variables I use would be meaningless without a lengthy introduction to them. You don’t need a deep math background to read this book though, it’s written in an approachable style that requires only 1st-year university math and some basic machine learning knowledge to understand, and I believe that is why this is the most popular textbook for beginners to RL.

I highly recommend this book if you are interested in what modern machine learning algorithms are capable of beyond just variations on linear regression and gradient descent. If you have the inclination, I also recommend completing the coding exercises in each chapter—coding up the algorithms helped me *feel* what was going on as if I were an RL agent myself, and this was the key that allowed me to draw the parallels between human and machine intelligence that I described in this review.

It took me four months to read this book and do the exercises, and I also did it as part of a class I am taking for my Master’s Degree, but it was well worth the investment. This book took me from barely any idea about reinforcement learning to being able to comfortably read the latest

RL research papers published on arXiv. Perhaps you'll also discover something new about AI that you didn't realize you were missing before.

You can [download the book for free](#) from the author's website.

Revolt of the Public by Martin Gurri

Gurri's *Revolt of the Public* offers an explanation for a lot of what makes the world seem so strange right now. Published in 2014, it foreshadows Trump's 2016 Presidential victory. But Gurri's thesis explains much more than just the Trump phenomenon. It applies equally to the Arab Spring of 2011 and the recent rise of Socialist ideas in the US. The simplified version of Gurri's thesis is as follows: The information age shows us the flaws of our leaders, leading to disillusionment and ultimately threatening democracy. But you really shouldn't shorten Gurri's thesis. He takes great pains to include nuance and qualify his arguments.

Here is my less simplified version of Gurri: He starts with the premise that the power of the elites is limited. Elites are the small group of people who run society through politics, business or other positions of power. To gain power or win elections, elites are forced to make big promises. They are destined to fall short because society is made of complex systems like the economy or the federal government. Complex systems are inherently unpredictable. For example, the chairman Federal Reserve cannot control the economy, no matter how brilliant she is. In the past, the media shielded elites from their failures. For example, the press gave JFK favorable coverage after the Bay of Pigs disaster. This was possible because media was dominated by a small number of institutions. Today, diverse information sources enable us to see the failures of our elites. This leads to disillusionment with government and our wider social system.

Before diving deeper, it is important to understand Gurri's definition of "the public." He is clear to say the public is not the people. The public is a group who have interest in an issue, but cannot directly influence it i.e. they are not elites. The public can unpredictably coalesce around specific issues, and disperse just as easily.

What has created this new information age? It isn't only the internet. It is all the diverse new sources of information. Cable news plays a big role. Gurri describes how Al Jazeera covered the Arab Spring when state-controlled media was censoring it. The internet is special for two reasons. First, it allows amateurs to provide an alternative to elite consensus. This can be a blogger opposing a dictator or a reddit poster peddling conspiracy theories. Second, social media enables interested groups to organize into a "public" around specific issues. Gurri profiles Wael Ghonim, a key figure in the Egyptian branch of the Arab Spring. Ghonim was not powerful or connected, but was able to spark a revolution using a Facebook group.

Gurri's examples come from a remarkably diverse set of groups. In just the US, he profiles groups from opposite ends of the political spectrum: Occupy Wall Street and the Tea Party. Other groups are products of completely different political contexts across the globe, from the Middle East to Western Europe. Reading about these protests in

real time, I had a sense they were somehow related, but I never heard anyone make a clear connection. Gurri makes a strong case they are part of the same “revolt of the public.” These groups share a few key traits that differentiate them from previous political movements. Each is rebelling against the elites. They don’t have strong ties to existing political parties or ideologies. For the most part, they don’t even pose a clear alternative – they just oppose the status quo. This last part is especially concerning to Gurri.

Gurri sees dark consequences from this loss of trust in elites. This is not some new golden age of participatory democracy. The revolts are about demolishing the existing system. The example groups do not pose credible alternatives. Most are unwilling to even name leaders.

This process begins to threaten democracy when politicians attack the system for personal gain. Gurri uses Obama as an example of a politician who ran against the establishment. His analysis of Obama feels a little forced to me. However, his analysis eerily describes Trump two years before he was elected. Trump’s run for president was an out-of-sample prediction test for Gurri’s hypothesis, and it passed with flying colors. He argues that a politician can leverage discontent with the system to get elected. Once elected, they will have no program or real ideology. All they can do is continue their opposition, even as they are in charge. This inevitably erodes trust in the fundamental institutions of democracy.

Gurri then suggests some potential solutions. The crude version of his solution is simply expecting less from government. Society, as a complex system, cannot be fully controlled by anyone. If we expect less from government, we won’t revolt in disappointment. The more nuanced version of his solution is about returning action to the personal sphere. The personal sphere consists of the everyday events that individual people can see and influence. The personal sphere is more manageable because it is not a complex system. Actions lead to predictable, or at least comprehensible results. To delegate to the personal sphere, government should increase transparency. Transparency will empower all individuals to relate policy to their personal sphere.

Revolt of the public is a great book. It pulls together disparate threads from the Arab Spring to the Tea Party. It analyzes trendy topics without getting caught up in hype or buzzwords. He avoids the trap of blindly blaming everything on social media. I appreciate how Gurri consistently highlights the uncertainties and limits of his analysis.

I don’t fully agree with his characterization of experts or government. I think he undervalues both. I won’t litigate that here. Even if he does underrate government or experts, it doesn’t damage his thesis. His thesis only requires experts and government to have limited powers. This is certainly true.

What about economic inequality? Numerous other analysts have blamed economic inequality for the trends described by Gurri. Inequality has a few things going for it as a

possible explanation. First, there is a clear causal mechanism: inequality increases the distance between the upper middle class and the elites. Second, the timing fits: Inequality has increased in the last few decades. Finally, inequality was a major grievance of several groups provided by Gurri, including Occupy Wall Street. I personally think inequality is a problem, but Gurri's work has convinced me it is not the main explanation here. Inequality fails to explain analogous right-wing movements like the Tea Party. Gurri's thesis is more powerful because it has explanatory power across the political spectrum. Our new information environment creates the conditions for a revolt of the public. Inequality is merely one issue the public will coalesce around when they revolt.

I found myself frequently contrasting Gurri's work with Ezra Klein's *Why We're Polarized*. Both authors offer compelling explanations for the disfunction in modern American politics. Klein identifies polarization as a key issue. Who is right? I see their approaches co-existing: Klein explains Mitch McConnell, while Gurri explains Trump and Bernie Sanders. In other words, Klein's polarization explains why traditional parties will not or cannot compromise to get things done. Gurri explains the emergence of radicals who reject the parties altogether. On the surface, these trends look similar. Trump certainly polarizes people. But Trump is not fundamentally a radical Republican. He is a radical against the system.

I'm not convinced by Gurri's proposed solutions. Maybe I didn't fully understand them. I also think this is a hard problem with no simple answers. Gurri deserves credit for articulating it. I believe he has identified an existential threat to modern democracy. To solve this, it is important to remember Gurri's lesson on complex systems: they are inherently unpredictable. I believe experimentation is the only path forward. We, as a society, will have to try a wide range of potential solutions to find the few that actually work.

