

# Book Review Runners Up

## Volume 2. S-W

[Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies by Geoffrey West](#)

[I](#)

[II](#)

[III](#)

[IV](#)

[V](#)

[VI](#)

[VII](#)

[Scientific Freedom: The Elixir of Civilization by Donald W Braben](#)

[The Model](#)

[Who Don Was Looking for: "The Planck Club"](#)

[The Cost of Venture Research](#)

[What Changed](#)

[Conclusion](#)

[Second Nature: A Gardener's Education by Michael Pollan](#)

[Shattered Sword: The Untold Story of the Battle of Midway by Jonathan Parshall and Anthony Tully](#)

[I: What is this book, and why am I reviewing it?](#)

[II: The gritty details](#)

[III: Analysis and Takeaways](#)

[Shut Out by Kevin Erdmann](#)

[I.](#)

[II.](#)

[III.](#)

[IV.](#)

[VI.](#)

[VII.](#)

[SPQR by Mary Beard](#)

[Steve Jobs by Walter Isaacson](#)

[I.](#)

[II.](#)

[III.](#)

[Strangers Drowning: Impossible Idealism, Drastic Choices, and the Urge to Help by Larissa MacFarquhar](#)

[I.](#)

[II.](#)

[III.](#)

[Strong Towns by Charles Marohn](#)

[Stubborn Attachments, by Tyler Cowen](#)

[In Search of A Free Lunch](#)

[Comparing Desires for Apples And Oranges](#)

[Time, and Other Illusions](#)

[The Common Sense Approach to Selling Your Baby](#)

[Butterflies and Hitler](#)

[Finishing the Philosophical Conversation](#)

[Ten Days That Shook the World by John Reed](#)

[I](#)

[II](#)

[III](#)

[IV](#)

[V](#)

[The Age of Entitlement: America Since the Sixties by Christopher Caldwell](#)

[I.](#)

[Unifying Force](#)

[The Worst Court Case in American History](#)

[It's All In the Game](#)

[The Soft Underbelly of the US Government](#)

[The New Constitution](#)

[Reconstruction 2.0 - The Largest Undertaking in the History of the American Government](#)

[The False Prophet](#)

[Pandora's Box](#)

[II.](#)

[The Costs of Entitlement](#)

[Racialization](#)

[White Malaise](#)

[Judicial Politicization](#)

[Financial Time Bomb](#)

[Coming Apart](#)

[III.](#)

[An Alcoholic Aside](#)

[IV.](#)

## Priors Adjustment

The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia by James C. Scott (2009)

Farming

Political structure

Ethnicity

Writing

Religion

The Book Of Why by Judea Pearl

The Burnout Society by Byung-Chul Han

I.

II.

III.

IV.

V.

VI.

VII.

VIII.

Epilogue

The Call of Cthulhu by H. P. Lovecraft

A Distinction in Mystery

Mysteries of Cthulhu

The Case Against Reality by Don Hoffman

I.

II.

III.

The Cult of Smart: How Our Broken Education System Perpetuates Social Injustice by Fredrik deBoer

The Culture of Narcissism by Christopher Lasch

The Eldraeverse by Alistair Young

The Great American Gamble - Deterrence Theory and Practice from the Cold War to the Twenty-First Century by Keith B Payne

Theory - Schelling

Theory - Kahn

Rationality?

Deterrence during the Cold War

Deterrence after the Cold War

Verdict

## [The Gulag Archipelago](#)

[Central theses](#)

[Resisting interrogation](#)

[Sentencing](#)

[Traitors](#)

[Wreckers](#)

[When everyone is a criminal, the cop is king](#)

[Labor camps](#)

[The Saga of Naftali Frankel](#)

[Channeling Corruption](#)

[More corruption](#)

[Women](#)

[Conservatives and kulaks](#)

[Changes 1950-1954](#)

[Factuality](#)

[Ending](#)

## [The Horse, the Wheel, and Language by David W. Anthony \(2007\)](#)

## [The Human Predicament: A Candid Guide to Life's Biggest Question by David Benatar](#)

[0. The Warning](#)

[I. The Asymmetry Between Pain and Pleasure](#)

[II. The Abysmal Quality of Life](#)

[III. The Meaninglessness of Existence](#)

[IV. The Certainty of Death](#)

[V. The Predicament](#)

[VI. The Objections](#)

[VII. The Solution](#)

[VIII. The Uncertain Future](#)

## [The Joy Luck Club by Amy Tan](#)

[Introduction: The Words You Cannot Say](#)

[Wounds: The Rifts Between Generations and Cultures](#)

[Interlude: Using the Category of Concrete-Relational Thinking to Decrypt the Narratives in "The Joy Luck Club"](#)

[Restored beyond Expectation: Possessors of Agency and of an Uncommon Vision](#)

[Wordplay, Structure, and, Imagery in "The Joy Luck Club"](#)

[Conclusion: Catharsis. Cost. Catalysis.](#)

## [The Millionaire Next Door by Thomas J. Stanley and William D. Danko \(and other books\)](#)

[You're Probably Wondering Why I Called You Here Today](#)

[The Millionaire Next Door](#)

[Your Money or Your Life](#)

[I Will Teach You To Be Rich](#)

## [So What Should I Read?](#)

### [The Ministry for the Future by Kim Stanley Robinson](#)

[Geoengineering](#)

[Terrorism](#)

[Financial Engineering](#)

[Something Boring: Nuclear Plants, Green Concrete, and Wheat Semi-Dwarves](#)

### [The No Breakfast Fallacy: Why the Club of Rome was wrong about us running out of resources by Tim Worstall](#)

[Keeping the Focus Narrow](#)

[The Fallaciousness of 'No Breakfast'?](#)

[Renewable Versus Finite](#)

[The Vastness Gambit](#)

[The Vastness Measured](#)

[A Triumvirate of Doubters](#)

[Surely You're Joking Mr Worstall](#)

[Abundance and Scarcity](#)

[More Abundance and Less Scarcity](#)

[Abundant Addenda](#)

### [The Open Society and its Enemies by Karl Popper](#)

### [The Oxford Handbook of 4E Cognition by Albert Newen, Leon de Bruin, and Shaun Gallagher](#)

### [The Power of the Powerless by Václav Havel](#)

[The "Post-Totalitarian" System](#)

[The Obedient Grocer](#)

[Ideology](#)

[Who Enforces the Ideology?](#)

[The Disobedient Grocer](#)

[Truth-Force](#)

[Living in the Truth as a Moral, not Political Revolution](#)

[Politics Under Post-Totalitarianism](#)

[Dissent and Opposition](#)

[Small-Scale Work](#)

[People and Politics](#)

[Revolt and Law](#)

[Building Parallel Structures](#)

[The Crisis of Contemporary Technological Society](#)

[What Is to Be Done?](#)

### [The Recovering by Leslie Jamison](#)

[The Recovering - The Fall](#)

[The Recovering - Whence Redemption?](#)

[The Current Landscape](#)

[GENDER](#)

[IDENTITY](#)

[Don't Invent the New, Remember the Old](#)

[An Identifying Audience](#)

[Why Does This Matter So Much?](#)

[The Red and the Black by Stendhal](#)

[The Righteous Mind by Jonathan Haidt](#)

[Part I: Intuitions Come First, Strategic Reasoning Second](#)

[Chapter 1: Where Does Morality Come From?](#)

[Chapter 2: The Intuitive Dog and its Rational Tail](#)

[Chapter 3: Elephants Rule](#)

[Chapter 4: Vote for Me \(Here's Why\)](#)

[Part II: There's More to Morality than Harm and Fairness](#)

[Chapter 5: Beyond WEIRD Morality](#)

[Chapter 6: Taste Buds of the Righteous Mind](#)

[Chapter 7: The Moral Foundations of Politics](#)

[The Care-Harm Foundation](#)

[The Fairness-Cheating Foundation](#)

[The Loyalty-Betrayal Foundation](#)

[The Authority-Subversion Foundation](#)

[The Sanctity-Degradation Foundation](#)

[Chapter 8: The Conservative Advantage](#)

[The Liberty-Oppression Foundation](#)

[Part III: Morality Binds and Blinds](#)

[Chapter 9: Why Are We So Groupish?](#)

[Chapter 10: The Hive Switch](#)

[Chapter 11: Religion is a Team Sport](#)

[Chapter 12: Can't We All Disagree More Constructively?](#)

[The Ruin of Kasch by Roberto Calasso](#)

[The Revolution](#)

[The Tale Told in the Desert](#)

[Sacrifice](#)

[Goethe Beholds the Paintings](#)

[The Young Hegelians](#)

[The Nihilists](#)

[The Fall](#)

[The Signal and the Noise: The Art and Science of Prediction by Nate Silver](#)

[The Silmarillion by J. R. R. Tolkien](#)

- [I.](#)
- [II.](#)
- [III.](#)
- [IV.](#)
- [V.](#)
- [VI.](#)
- [VII.](#)
- [VIII.](#)
- [IX.](#)
- [X.](#)
- [XI.](#)
- [XII.](#)
- [XIII.](#)
- [XIV.](#)

[The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron by Bethany McLean and Peter Elkind](#)

- [I.](#)
- [II.](#)
- [III.](#)
- [IV.](#)
- [V.](#)
- [VI.](#)

[The Sovereign Individual by William Rees-Mogg and James Dale Davidson](#)

- [History](#)
- [Nation-states](#)
- [Government as a service](#)
- [Inequality](#)
- [The next steps](#)
- [Random passages I liked](#)
- [Other predictions](#)
  - [First, stuff they got wrong.](#)
  - [Next, stuff they got right.](#)
- [What to do about it](#)

[The Sovereign State and Its Competitors by Hendrik Spruyt \(1994\)](#)

- [France](#)
- [Germany](#)
- [Italy](#)

[The Suicide of Reason by Lee Harris](#)

- [I. The Alien Space Ship](#)
- [II. So, where did the process come from?](#)

[III. The Other Gods](#)

[IV. The Crash of Civilization](#)

[V. How screwed are we?](#)

[The True Believer: Thoughts on the Nature of Mass Movements by Eric Hoffer](#)

[I.](#)

[II.](#)

[III.](#)

[IV.](#)

[The True Creator of Everything: How the Human Brain Shaped the Universe as We Know It by Miguel Nicolelis](#)

[A Totally Organic Experience! \(With a Tip of the Hat to Clair\)](#)

[Check out the Big Brainet on Brad!](#)

[What is Mind? No Matter. What is Matter? Never Mind](#)

[The WEIRDest People in the World by Joseph Henrich](#)

[I. Systemic Evolution](#)

[II. Biological-Cultural Fit](#)

[III. Apes \(Pre-200,000 BCE\)](#)

[IV. Early Homo Sapiens \(200,000 BCE - 10,000 BCE\)](#)

[V. Kin-Based Structures in Agrarian Age \(10,000 BCE - 500 CE\)](#)

[Patrilineal Descent](#)

[VI. Catholic Church in the Middle Ages \(500 CE - 1500 CE\)](#)

[VI.A How the Church's Marriage and Family Program \(MFP\) Broke Kin-Based Institutions](#)

[VI.B The Creation of Non-Kin-Based Voluntary Organizations](#)

[VII. Protestant Reformation and the Industrial Revolution \(1500-2000\)](#)

[VII.A Doux Commerce as Domesticated Intergroup Competition](#)

[VII.B Individual Rights and Democracy](#)

[VII.C Individualistic Science and the Collective Brain](#)

[VIII. Conclusion](#)

[The WEIRDEST People in the World by Joseph Henrich](#)

[The World Imagined by Hendrik Spruyt \(2020\)](#)

[Chinese Tributary System](#)

[Islamic Cultural Community](#)

[Galactic Empires of Southeast Asia](#)

[Things Fall Apart by Chinua Achebe](#)

[Introduction](#)

[Characters and Plot](#)

[Life in Umuofia](#)

[Language](#)



[Spirit Masquerades](#)

[Infant Mortality](#)

[Miscellany](#)

[The Rise of Christianity](#)

[A Metaphysical Reading](#)

[Questions](#)

[Is Things Fall Apart an Igbo nationalist work?](#)

[How would Achebe have fixed the problems with the traditional ways?](#)

[How had Igbo society already changed from past contact?](#)

[Conclusions](#)

[Transcendence: How Humans Evolved through Fire, Language, Beauty, and Time by Gaia Vince \(2020\)](#)

[Trustee report review](#)

[Fletcher's Investors](#)

[Fletcher's Compensation](#)

[Cashless Exercise](#)

[An Academic Interlude](#)

[Taking Stock](#)

[Understand \(1991\) by Ted Chiang](#)

[shaman / superman](#)

[accidental / chosen](#)

[his pre-K self is predictive of his post-  
the Story gets better. did Chiang get better?](#)

[a trite metaphor for Connection](#)

[excessive growth of the optic nerve](#)

[the Misunderstood Genius](#)

[the Brain is amped; and with it, the Body](#)

[an unreliable Narrator? How to do a Super-mind, first-person?](#)

[The Fifth Ampule? all he needs is inward/outward unity](#)

[good guy / bad guy](#)

[how do gods speak?](#)

[how do gods fight? Clarke's 3rd Law](#)

[Does Greco get in death what he'd sought?](#)

[the Self-destruct command: Chiang as thy Destroyer](#)

[the smartest in the Zoo](#)

[addressivity](#)

[The Fifth Ampule, reconsidered](#)

[final pitch: the Pantheos](#)

[Walkaway by Cory Doctorow](#)

!

[II](#)  
[III](#)  
[IV](#)  
[V](#)  
[VI](#)  
[VII](#)  
[VIII](#)

#### [What the Buddha Taught by Walpola Rahula Thero](#)

[Introduction](#)  
[As Introduction to More Buddhism](#)  
[Almost-Religion](#)  
[The Poisoned Arrow](#)  
[The Four Noble Truths](#)  
[Suffering and the Self](#)  
[Metaphysics](#)  
[Meditation](#)  
[Conclusion](#)  
[Notes](#)

#### [Where Is My Flying Car? A Memoir of Future Past by J. Storrs Hall](#)

[I](#)  
[First we learn to fly our cars](#)  
[Then we learn about limitless power](#)  
[And moving on to the beautiful promise of Cold Fusion](#)  
[And we build palaces, atom by atom](#)  
[Building the Tower of Babel and 100km space piers](#)  
[And a consequent dilution of culture](#)  
[II](#)  
[Onerous regulations](#)  
[Failure of nerve and imagination](#)  
[III](#)  
[Coda](#)

#### ['Without Gloves', 'Without Grease', 'The Great Game of Politics' and 'Political Behavior' by Frank R Kent](#)

#### [Working With Contracts: What Law School Doesn't Teach You by Charles M. Fox](#)

[Background Knowledge](#)  
[What Do Contracts Do?](#)  
[Parallels to Software Development](#)  
[The "Do What I Mean" Button](#)  
[Materiality](#)  
[Reasonableness](#)

[Behaving Normally](#)  
[Remedial Provisions](#)  
[Takeaways](#)

# Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies by Geoffrey West

*Scale* (2017) is a general audience (for the “intelligent layperson”) book which relates the author’s - a physicist at the Santa Fe Institute - research into the life sciences and later into cities, and as such is hard to categorize under a single topic or thesis. We could look to the title, which refers to scaling laws or power laws. That is, relationships of the form  $y = k(x^s)$  as opposed to linear relationships  $y = mx + b$ . Or we could look to the subtitle, *The Universal Laws of Life, Growth, and Death in Organisms, Cities, and Companies*. My own irreverent-but-not-irreverent characterization is that, in addition to describing a certain branch of science called complexity theory, this is a book which argues biologists, economists, sociologists, etc. should think more like physicists. The dominant thread in *Scale* involves simple mathematical rules showing up in unexpected places – a hidden order where one would expect lots of statistical noise. With respect to certain properties all mammals are scaled versions of each other, as are all birds, as are the cities of a given country, as are companies. It’s a bold claim, particularly the latter two items (biology experts may already know about scaling between animal species), but West backs it up.

A heads up – many of the books reviewed on this blog have a single, well-defined thesis that they argue for the entire book. *Scale* is not that. This is partly because it is a layperson-level summary of research, and partly because it moves around a lot. In some ways it is one of those books like *Godel, Escher, Bach* where you end up with a mix of things that all happen to be in the author’s head, and you can understand how they’re connected if you read it but trying to explain it concisely is a challenge. I am writing this review mostly to share some of the phenomena described in it, but there are a couple of debatable ideas and I will comment on those.

|

Before getting into the core topics of the book, West provides several samples and preliminaries. To highlight one, consider drugs. If you know how much of a drug constitutes a dose for an adult human, how much should you give to something that’s larger or smaller than the adult human? There is no simple answer to this question, but

we can talk about estimates. If you were forced to come up with a rough guess for how to change the dosage, you would probably assume that drug dosage is proportional to body mass, i.e. if a 6 lb baby should be given 40 mg doses of Infants' Tylenol, then a baby weighing 36 lb should be given doses which are 240 mg – weight multiplied by 6 means dosage multiplied by 6 as well. Of course, this is the kind of simple question that only gets brought up if the natural response happens to be wrong. And indeed,

“However, regardless of details, an understanding of the underlying mechanism by which drugs are transported and absorbed into specific organs and tissues needs to be considered in order to obtain a credible estimate. Among the many factors involved, metabolic rate plays an important role. Drugs, like metabolites and oxygen, are typically transported across surface membranes, sometimes via diffusion and sometimes through network systems. As a result, the dose-determining factor is to a significant degree constrained by the scaling of surface areas rather than the total volume or weight of an organism, and these scale nonlinearly with weight.”

If the length/height of a 3D entity increases by a factor of  $c$ , then surface areas increase by  $c^2$  and volume by  $c^3$ , assuming the thing stays proportional to what it was before. The relationship between size and dosage of various drugs is not as clean as some of the other results in the book, so.. don't try this at home. But, if you absolutely had to estimate the dose of a drug, and you knew the appropriate dose for someone (or some animal) of a given size, you could use a scaling law with an exponent of  $2/3$ . If you've calibrated baby Tylenol dosage for a 6 lb baby and you know that the correct dose is 40 mg, then for a baby which is 6 times as large the dose should be  $6^{(2/3)}$  times as large – this comes out to about 132 mg, around half of the uninformed estimate of 240 mg. And to reiterate, don't try this at home, because it's meant to be a familiar example to get across the general idea of scaling laws rather than medical advice, and the scaling law here isn't that accurate anyway. But, this example is not made up. West pulled those numbers off the bottle. He notes that the company's website “now wisely [recommends] that a physician be consulted for babies less than 36 pounds”, but that “Nevertheless, other reputable Web sites still recommend linear scaling for babies younger than this”.

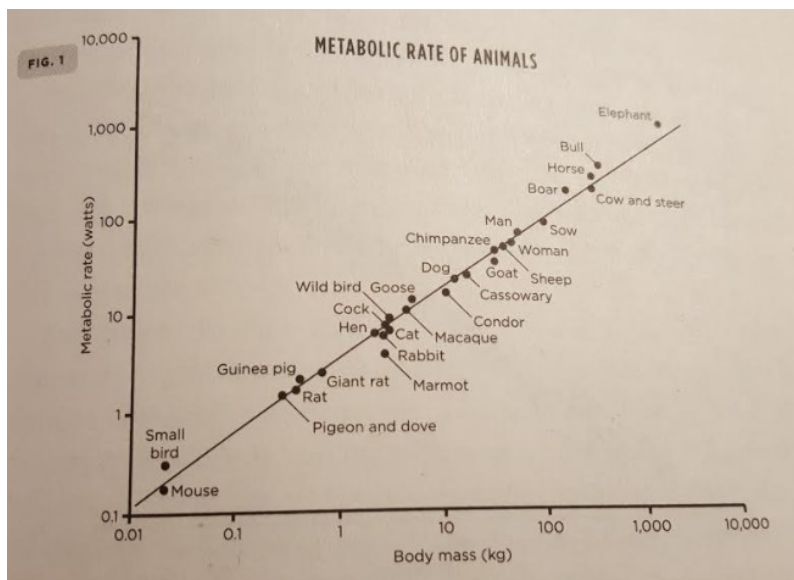
Perhaps there is some hidden reason for the linear scaling of Infant Tylenol dosage. But West provides another story of failure to understand how doses should scale. In 1962 two psychiatrists and a zoologist attempted to study the effects of LSD on elephants. It was not clear, however, how much LSD should be injected into an elephant. The solution they came up with was to take a dose suitable for a cat and assume that a dose should be proportional to body weight - a linear relationship. I filled in some math based on the average weight of each animal and came up with an estimate that this is 8 to 12 times more than what West says their estimate ought to have been. The results: “Poor

old Tusko died an hour and a half later. Perhaps almost as disturbing as this awful outcome was that the investigators concluded that elephants are ‘proportionally very sensitive to LSD’.” This study was published in the elite journal *Science*.

## II

So, if you’re looking for how two quantities relate, the impulse is to assume a simple ratio  $y = k \cdot x$ . But instead it might look like  $y = k \cdot (x^s)$ . What can we describe with  $y = k \cdot (x^s)$ ? Most of what’s in the rest of the book.

Start with animals. What is the metabolic rate (how many calories does it eat each day) of the average member of a given species? You would probably expect a given cell to



take the same amount of energy for maintenance regardless of how big the animal is, and thus be proportional to the mass of the animal. Alternatively, through the power of context and leading questions you’ve guessed that metabolic rate scales like a power law with an exponent other than 1. This would be correct. The exponent is  $3/4$ . What about heartbeat, or life expectancy?

The number of heartbeats across an entire lifetime is roughly invariant across any taxonomic group – every mammal from rats to dogs to whales has about  $10^9$  heartbeats in the course of its lifetime (humans have far more but that is due to medical technology). Heart *rate*, however, decreases systematically with increasing size, with an exponent of about  $-1/4$ . There is a corresponding increase in life span, and in age until maturity. The ratio of white and grey matter in the brain, too, fits into a scaling law with an exponent of about  $3/4$ . What’s even more striking here is that the models are quite accurate. The biologists call them *allometric scaling laws*. West says that biologists have known about scaling since before 1900, and that the pattern was well-documented by Max Kleiber in 1932. But with a couple exceptions, they by and large did not ask the right questions:

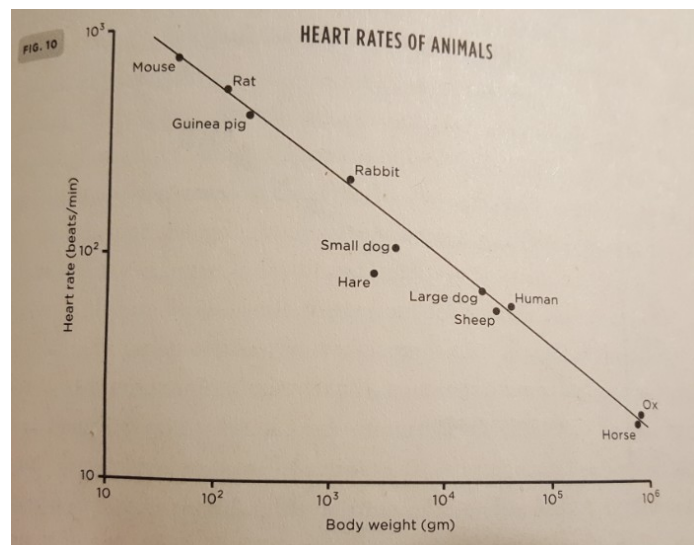
“In the 1980s several excellent books were written by mainstream biologists summarizing the extensive literature on allometry. Data across all scales and all forms of life were compiled and analyzed and it was unanimously

concluded that quarter-power scaling was a pervasive feature of biology. However, there was surprisingly little theoretical or conceptual discussion, and no general explanation was given for why there should be such systematic laws, where they came from, or how they related to Darwinian natural selection.”

“As a physicist, it seemed to me that these ‘universal’ quarter-power scaling laws were telling us something fundamental about the dynamics, structure, and organization of life. Their existence strongly suggested that generic underlying dynamical processes that transcend individual species were at work constraining evolution. This therefore opened a possible window onto underlying emergent laws of biology and led to the conjecture that the generic coarse-grained behavior of living systems obeys *quantifiable* laws that capture their essential features.”

There’s a criticism of biologists here and ... unless West happens to be outright wrong about the facts it seems to be well-founded. Can a biologist or someone from a related field weigh in here? How does any researcher look at that kind of data and not ask why? The book quotes a Nobel Prize winner from biology, Sydney Brenner, who said roughly the same thing about his own field: “drowning in a sea of data and thirsting for a theoretical framework with which to understand it...”

I’ve probably given you all an impression that Geoffrey West has an ax to grind against biologists. He doesn’t, sort of – he tells us a younger version of him did have such an ax but it’s a friendly thing now. What he does do stress a prescription for research, that it should be more multi-disciplinary and not conducted in silos. This too seems quite reasonable? Clearly there is a need for researchers to spend a long time figuring out highly specific and technical answers to highly specific and technical questions, but it doesn’t do a lot of good if nobody puts together a broader picture. Moreover, it’s ideal if the broader picture is something that more than 200 extremely specific people in the world can actually understand and follow. Besides the benefit of diverse perspectives/methodologies/yada yada, I’m willing to bet that having more multidisciplinary research would help with the dispersion of ideas. Maybe with more communication between fields, pharmaceutical companies wouldn’t have made an apparently elementary error like assuming that medicine dosage should be proportional



to mass instead of surface area. There is something of a motif in *Scale* where Geoffrey West recounts stories of professionals or researchers getting something badly wrong, usually because they missed something that was basic knowledge to another field. It's a theme which I've chosen not to focus on for this review, but examples include the Millennium Bridge falling victim to resonance, BMI being a patchwork of linear relations instead of using a scaling law, and the discovery of fractals' ubiquity being delayed for 15 years because the original discoverer (Lewis Richardson) buried it in an appendix to an unrelated article.

### III

So, in the mid 90s Geoffrey West sees scaling in animals and because he's a physicist he wants to craft a grand theory that explains why this happens. He teams up with biologists James Brown and Brian Enquist, and they give each other a crash course in their respective fields and methodologies. There are some culture shocks, etc., and over the course of the next few years they discover and publish in articles their theory about why there are scaling laws among taxonomic groups. The book goes on to outline that theory.

Repeating the quoted passage from earlier, let's make note of a truism. Evolution optimizes. Metabolic rate is the amount of energy a creature has to imbibe to maintain itself and continue living. If there was a way for a given species to become more energy efficient, to eat less while not sacrificing fitness in other areas, we would expect evolution to find it. The scaling for metabolic rate with respect to size, and its accuracy, strongly suggests that there is some physical law which at some point constrains how energy efficient an animal of a given size can be. The fact that so many of the exponents are close to fractions with a denominator of 4 strongly suggests that the number 4 has some fundamental importance here. What might that law and that importance be? To answer this we need to talk about fractal geometry. This is probably familiar, but here's a quick review:

For most people the word "geometry" brings to mind Euclidean geometry. This is the world of straight lines, perfect circles, right angles, etc. Or, more broadly, it is the world of smooth surfaces which can be approximated by straight lines should one zoom in far enough. However, while this paradigm does an excellent job of measuring and describing objects built by humans, it is decidedly less excellent at measuring and describing a number of objects in nature, which are recursive or which have detail "on many scales".

The canonical example, dating back to Mandelbrot's 1967 paper, is a coastline. Look at a coast, say that of Norway, on a world map. It is best described as "crinkly". Look at a map that is of Norway alone and it is still crinkly; the zoomed-in map captures features which the world map did not. Zoom in further and it will exhibit yet more features not





visible on the maps with a larger scope and lower resolution. In the real world we can't do this indefinitely; at some point we reach the resolution of atoms. But there is a broad range of scales at which we could look at the coastline and see interesting patterns. The same applies to broccoli, to river deltas, and to stock price charts: a stem of broccoli is rather like a

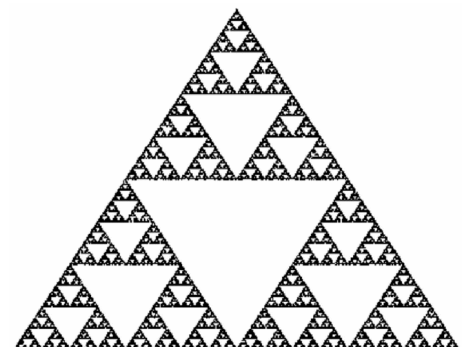
smaller copy of the entire head, financial charts on the scale of minutes or of weeks look about the same if you remove the numbers from the axis, and if you zoom in on a picture of a delta you end up looking at something similar to what you were looking at before.

The connection between fractals and power laws is that the former is one source for the latter. To understand this, think back for a moment to the Euclidean world. Suppose one doubles every side of a square. The area of the square itself, a 2-dimensional object, is multiplied by  $4 = 2^2$ . Double every side of a cube, a 3-dimensional object, and the volume is multiplied by  $8 = 2^3$ . If we allow this

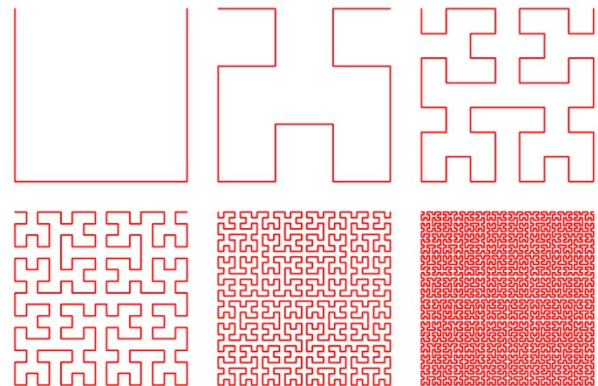
observation to serve not as a consequence but rather as the definition (or, to be pedantic, the motivation for an expanded definition) of dimension, then we can have dimensions that are numbers other than 1,2,3,...

. The fractal dimension is how much bigger the object gets if you scale it up. Or for the real-world examples, fractal dimension tells you how much longer/larger you're going to measure for a quantity if you use a more fine-grained measurement and capture more of the crinkles with your measuring stick or equivalent.

This brings us to a counterintuitive idea, that there can be a distinction between the dimension in the sense of taking  $n$  parameters to describe a location within the object, and dimension in the sense of how much bigger the object gets if you scale lengths. Or to go with a less correct but more intuitive explanation, imagine a flat two-dimensional sheet that you have made so crinkly, on scales going all the way down, that it manages to have something like volume. The bizarre mathematical construct that is Hilbert's space-filling curve turns out to have real-world relevance.



The type of fractal that matters here is a branching network, i.e. the circulatory system. Blood starts out pumped by the heart, and goes out through the aorta. The aorta then branches off into arteries. Those arteries branch off into other arteries, which eventually branch into arterioles and then finally into capillaries, which must reach every cell. Then the same thing happens in reverse, from bottom up. Nature's fractals aren't infinitely detailed, but if we just look at the first couple steps after the heart, the system doesn't much change each time we branch off into a smaller vessel. And if we compare animals of two different sizes, the main difference between their circulatory systems is the number of branching steps it takes before bottoming out into capillaries.



In terms of length the circulatory system is very, very large - about 60,000 miles for an adult human. But from the fractal viewpoint, length isn't quite the correct perspective to take. Instead we need to think about the fractal dimension. Remember Hilbert's space-filling curve? Blood vessels exist in three-dimensional space. But they are so detailed and branching that they can be thought of as having almost an extra dimension to them, through the power of fractal geometry. Disclosure: West's actual analogy doesn't involve Hilbert, but rather two-dimensional clothes in a three-dimensional washing machine. Either way, I have to admit that the analogy doesn't 100% make sense to me – Hilbert's line or crumpled clothes in the wash actually exist in the extra dimension they're filling, and with blood vessels the fourth dimension is more like a metaphor rather than something they literally fold branch into and fill. I assume West is sparing us a much more complicated mathematical analysis. The basic idea that you get more use out of space if you have a highly detailed, intricate network is simple enough, but why the scaling exponent should be exactly 4 – key for being the number of dimensions we live in plus 1, according to the book – is not something I feel comfortable expositing. If you're sufficiently curious, look up his team's paper, [The Fourth Dimension of Life \(1999\)](#).

The fractal-based theory ends up doing a great job of predicting several quantities related to blood vessels, as well as the analogous structures in flora. It also does a good job of predicting those big-picture quantities mentioned earlier, like growth curves and metabolic expenditure per cell. Hearts across all mammals maintain the same blood flow – speed and pressure. If we scale up an abstracted mammal, the circulatory system experiences predictable gains in efficiency. A horse's heart does not have to work as hard in an hour as does a rat's, although over the course of their respective lifetimes both hearts do around the same amount of work.

The theory can do even more. It can predict the size of the smallest mammals. The key to this prediction is that large blood vessels and very small blood vessels – capillaries – do not function in the same way. Your arteries have a pulse emanating from the heart. But after they branch enough times and become sufficiently narrow the pulsatile system isn't efficient; instead blood flow is steady at the lowest levels. West compares it to DC vs AC electrical current. And crucially, the tipping point between pulsatile and steady doesn't depend on the size of the animal; all that the size affects is how many steps there are before we reach the tipping point. Another way to put this is that arteries are fractal but capillaries are not. Shrews are among the smallest mammals; they only have one or two layers of artery before bottoming out into capillaries. A mammal that was any smaller would have a design based on a heartbeat but would have no pulse because its entire circulatory system would be based on steady flow. Such a design would be inefficient, and so it doesn't occur.

Briefly: West and co. are also able to say something about upper limits to size using the network viewpoint. As an animal gets larger, the capillaries very slowly (exponent  $1/12$ ) become further apart from one another. If they spread too far apart, they don't actually supply every cell and this hypothetical animal is interspersed throughout with cells which have no oxygen. Now, for land animals we're probably best off to ignore this and instead just make note of the square cube law – that the strength of limbs scales like a square and the mass of the things they're supposed to be holding up scale like a cube – but the network theory is able to predict something near the size of the blue whale as the largest possible mammal.

## IV

The first half of *Scale* talks about an aspect of biology which was already known, but not widely cared about or well-understood, and it both popularizes what came before and advertises the author's own contributions. Certainly biological allometry was not known to me, but since I have not taken biology since high school that is not too surprising. (Although, if the various pre-med students I knew in college learned this fact and never brought it up at gatherings, I'm disappointed, because I think it's pretty cool.) The second half of the book leaves biology behind, except as an analogy. West uses a metaphor – that of the city as a living entity – to motivate research. After all, scaling laws in biology came down to rules governing how infrastructure scales as a body gets larger, and cities have infrastructure too - pipes and roads and electrical wiring instead of blood vessels and the respiratory system. There was a researcher in Germany named Dirk Helbing who had looked at scaling in cities back in 2004; Geoffrey West recruited him and others into a collaboration to revisit and expand upon these ideas. The fact that there is a second half of the book suggests that the findings are going to be interesting, and indeed they are. Just as there were several biological quantities

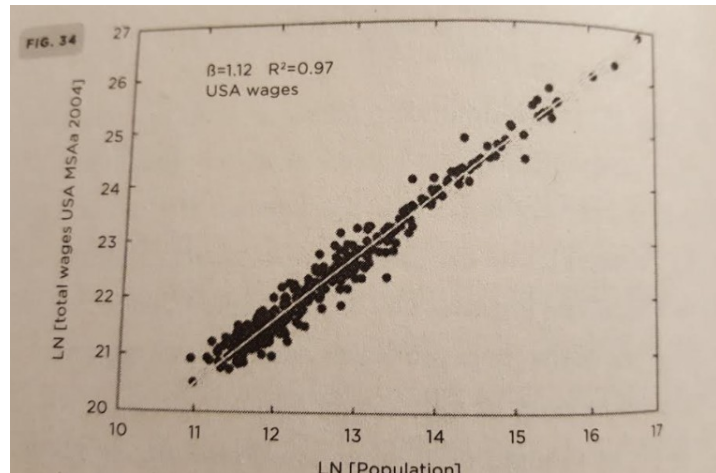
which one could predict for any mammal knowing only its mass, there are several quantities pertaining to cities which one can predict knowing only the population and the country in which it's located. (The models only compare the cities of one country to each other. It turns out that for this set of  $y = k \cdot (x^s)$ ,  $k$  depends on the country you're in and  $s$  does not, making country rather like taxonomic group in the analogy)

“Regardless of the specific urban system, whether Japan, the United States, or Portugal, and regardless of the specific metric whether the number of gas stations, the total length of pipes, roads, or electrical wires, only about 85 percent more material infrastructure is needed with every doubling of city size. Thus a city of 10 million people typically needs 15 percent less than of the same infrastructure compared with two cities of 5 million each, leading to significant savings in materials and energy use.”

“This savings leads to a significant decrease in the production of emissions and pollution. Consequently, the greater efficiency that comes with size has the nonintuitive but very important consequence that on average the bigger the city, the greener it is and the smaller its per capita carbon footprint. In this sense, New York is the greenest city in the United States, whereas Santa Fe, where I live, is one of the more profligate ones. On average, each of us in Santa Fe is putting almost twice as much carbon into the atmosphere as New York. This should not be thought of as somehow reflecting the greater wisdom of New York's planners and politicians, nor as the fault of Santa Fe's leadership, but rather as an almost inevitable by-product of the dynamics underlying economies of scale that transcend the individuality of cities as their size increases. These gains are mostly unplanned, though policy makers in cities can certainly play a powerful role in facilitating the enhancing the hidden 'natural' processes that are at work. In fact, this is a large part of what their job is. Some cities are very successful at doing this, while others are much less so. ...”

“These results are very encouraging and provide powerful evidence in support of the quest for a possible theory of cities. However, of even greater significance was the surprising discovery that the data also reveal that socioeconomic quantities with no analog in biology such as average wages, the number of professional people, the number of patents produced, the amount of crime, the number of restaurants, and the gross urban domestic product (GDP) also scale in a surprisingly regular and systematic fashion ...”

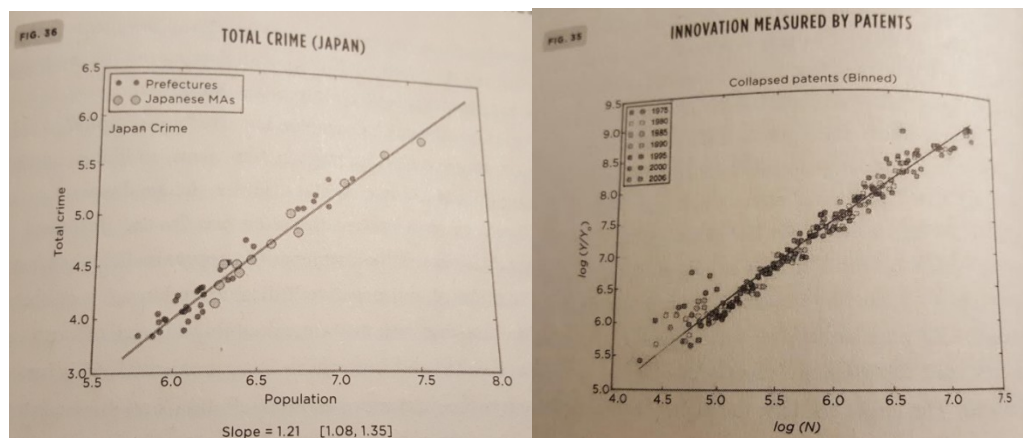
“Also clearly manifested in these graphs is the equally surprising result that all the slopes of these various quantities have approximately the same value, clustering around 1.15. Thus these metrics not only scale in an extremely simple fashion following classic power law behavior, but they all do it in approximately the same way with a similar exponent of approximately 1.15 regardless of the urban system.”



(The wages aren’t adjusted in any way for standard of living, but it is still noteworthy, and perhaps there is some insight about inflation lurking here. Or perhaps it is telling us about the concentration of the wealthy; West does use total wages rather than take a median.)

Later on:

“These results are pretty amazing. We typically think of each city, and especially the one we live in, as being unique ... Boston not only looks different but also ‘feels’ different from New York, San Francisco, or Cleveland, just as Munich looks and feels different from Berlin, Frankfurt, or Aachen. And they do and are. But who would believe that within their own urban systems they are approximately scaled versions of one another, at least as far as almost anything that you can measure about them is concerned? If you are given the size of a city in the United States, for example, then you can predict with 80 to 90 percent accuracy what the average wage is, how many patents it’s produced, how long all of its roads are, how many AIDS cases it’s had, how much violent crime was committed, how many restaurants there are, how many doctors and lawyers it has, et cetera. Much about a city is determined simply by its size.”



*Log of population against log of crimes, Japanese cities*

That larger cities have more criminals and more inventions per capita is in itself not at all surprising. What's incredible here is the accuracy. The figure plotting the logarithm of total wages in a city vs. the logarithm of population is a line with R-squared equals 0.97! I will admit to picking this one because the fit is especially good, and not all of the plots have an R-squared printed with them, but based on eyeball analysis this model is excellent even before one takes into account its extreme simplicity.

So why does this model work, and work better than it has any right to? "A faster pace of life" perhaps, or by analogy perhaps increasing the size of a city as raising the temperature of a gas and so increasing the number of collisions between the particles. In truth the question is not answered in a detailed or mathematically rigorous fashion, though not for lack of interest. The most promising direction for a rigorous explanation appears to be the rate at which interactions between humans happen. West tells us about phone call data: The amount of time spent on reciprocated phone calls scale with population, with an exponent of about 1.15. And that's reciprocated phone calls, meaning that both people called each other within some time frame. If one includes the non-reciprocated phone calls such as sales pitches the exponent is even higher; West doesn't say by how much. For those curious, this particular study was done in 2014. It's definitely plausible, but not completely convincing. West's team explores a natural hypothesis: We know that the maximum number of contacts a person can have, with varying degrees of closeness, is a human constant that doesn't depend on culture or the number of people in their settlement – see Dunbar numbers. But what about the number of your friends and acquaintances who know other friends and acquaintances of yours? Perhaps it is lower if you live in a city, because you have more spheres and so less chance for two of your friends to know each other? But, no. The team doesn't find any individual-level differences in social circles that change with population density. All the same, my priors tell me that there must be something here, and West seems to agree. He speculates about cities being diverse because they are large, and that this

helps people be able to find others who are very similar to them. My own nebulous thoughts involve the idea that population-dense areas are good if you want to do business (in a general sense of the word) but not if you want to, say, raise children. I'm sure an entire book could be written summarizing different hypotheses and the status of the research here, but I'm hoping that if a solid answer was known that West would have heard about it and mentioned it.

(Also, I have to wonder how the Internet age affects all of this. You can use Zoom from just about anywhere in the developed world, and the quality of the communication is the about the same, regardless of how dense of an area you're in.)

## V

Moving on from the inconclusive discussion that comes from "Why?", let's consider something else. If doubling the size of a city leads predictably to a 15% increase in the GDP per capita and the crime rate, and a 15% savings in the per capita cost of infrastructure networks such as roads, water systems, etc. then per capita measures are a poor benchmark for comparisons between cities. New York City would be at or close to the superlative in everything, simply for being the biggest city in the American urban system. And if the simple model predicts socioeconomic quantities with 80-90 percent accuracy, is there anything that can be said about the other 10-20 percent? Both of these questions suggest that one adjust for population, but using the power law model instead of per capita calculations to do so. West and his team did this on a set of 360 American cities:

"Amusingly, from this point of view, New York City as a whole turns out to be quite an average city, marginally richer than its size might predict (rank 88th income, 184th in GDP), not very inventive (178th in patents), but surprisingly safe (267th in violent crime). On the other hand, San Francisco is the most exceptional large city, being rich (11th in income), creative (19th in patents), and fairly safe (181st in violent crime). The truly exceptional cities are typically smaller ... "

"This is just for a single year (2003), and it is natural to ask how any of this changes with time. Unfortunately, readily accessible data on all of these metrics are hard to come by prior to about 1960. However an analysis covering data over the last forty to fifty years reveals some intriguing results ... Cities that were overperforming in the 1960s, such as Bridgeport and San Jose, tend to remain rich and innovative today, whereas cities that were underperforming in the 1960s, such as Brownsville, are still near the bottom of the rankings. So even as population has increased and the overall GDP and standard of living have



risen across the entire urban system, relative individual performance hasn't changed much."

West notes that San Jose is the location of Silicon Valley, so rather than big tech explaining why San Jose punches above its weight, it is more likely that San Jose punching above its weight helps explain why Silicon Valley is where it is.

(As an aside, smaller entities being more exceptional is to be expected, is it not? The smaller the sample, the larger a deviation from the mean is plausible. Even a small county is a large sample, so maybe not, but 1. We are given only the relative rankings; no other metric is provided. And 2. If we view cities as a set of neighborhoods, rather than people (which seems reasonable, if what we're interested in here is income, crime, and patents; neighborhoods de facto segregate with respect to the first two), then the size of each sample is not so large. The book says nothing more on how size affects the results than what I quoted above, so factor in high epistemological uncertainty to this assessment. But, I would be curious to see if increases (decreases) in relative population could predict movements toward (away from) the mean rankings of GDP and crime.)

This is a general audience book on science rather than history or sociology, and the emphasis is on informing rather than persuading. West occasionally delves into politics, mostly to advocate for sustainability, environmentalism, and acknowledgement of the dangers of exponential growth. But by my view his team's findings about cities have powerful implications for political theory. This research is describing a strong and fundamental relationship between geography and productivity. We've all heard over the past few years the story that "big urban centers are the winners of globalization and low-population rural areas are missing out". But this research is suggesting that it does not matter what any specific government does (short of abject totalitarianism – see N. Korea). Not only will bigger settlements always have higher crime and production per capita, it will be in amounts that are easy to predict with high accuracy. There's room for "culture" to matter, since Japanese cities are distinct from American cities by these numbers. But good luck changing culture if you're a mayor or a city council trying to improve your municipality.

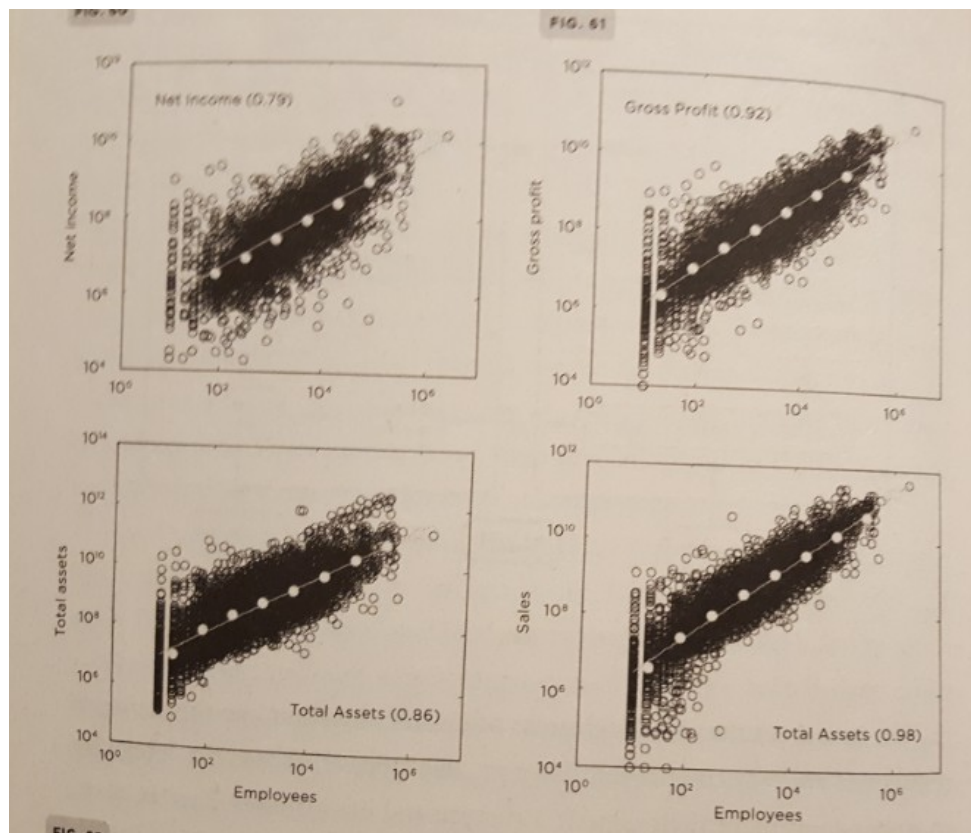
As you might recall from one of the above passages, West himself disagrees and says that if policy makers are smart about the natural processes and take a view of decades instead of election cycles they can achieve great things. He spends pages elsewhere criticizing top-down planned cities like Brasilia, but evidently believes there are other policies out there that could guide a city constructively. I hope he's right, but he doesn't say anything about what these policies might look like, and based on the evidence this section focuses on I draw a different conclusion. If the over (under) -performing cities are the same as they were 50 years ago, then why would we expect that the policies of one city or another made any difference? Or an even better question, why would we expect a city's politics to exist outside of the powerful generic forces being described here? There may well be aspects of city hall's priorities that can be predicted based on



the size of the city; after reading this book it wouldn't surprise me. (What I'm trying to say here is, maybe it's like [Does Reality Drive Straight Lines On Graphs, Or Do Straight Lines On Graphs Drive Reality?](#) but with  $y = k \cdot (x^s)$  instead of  $y = mx + b$ .) West seems to take as his null hypothesis that the root cause for the differences in ranking is smarter policy. I look at the inflexibility in relative advantages and disadvantages over the course of a half century, and my own null hypothesis is that the root cause is nationwide self-sorting by class. On the other hand, maybe Geoffrey West just didn't want to be controversial. And from a certain point of view, it is possible for smart and long-term-minded mayors and council members to be voted in and enact awesome policies that cut down on crime and generate wealth. It just never happens, because "socially/economically possible" is a much smaller space of events than "possible possible". (see also [WebMD And The Tragedy of Legible Expertise](#))

## VI

Following the section on cities is a short section on companies. This time the data is very noisy, with lots of spread, but there are still observations to make. This will be brief because the theory is not nearly as well-explored as the biology and the study of cities, but:



*Horizontal axis is log of number of employees in all cases; vertical axes are net income, gross profit, total assets, and sales. Data set is all 28,853 public companies in the United States between 1950 and 2009*

The odds that a company will disappear within a given time frame does not depend on how big it is. Nor how old (up to a point, back to this in a sec)

The fact that the chance that a company “dies” (is bought out or goes bankrupt) without regards to size means that we can meaningfully speak of a half life for companies; for publicly traded American companies it is currently around 10.5 years.

The rate of turnover of the Fortune 500 has increased. In 1958 the average lifetime of a spot on the Fortune 500 was measured 61 years; in 2017 it was 18 years. Note that falling off the list does not always mean bankruptcy or merger. It could just mean contraction.

The amount that a company spends on R&D as it increases in size decreases a proportion of the overall budget. This suggests (though does not show definitively) that administrative overhead outpaces adaptability as a company grows - no surprises there. A purely half-life decay theory would tell us that there are basically no companies older than 200 years. This is not true. From the book: “According to the Bank of Korea, of the 5,586 companies that were more than two hundred years old in 2008, over half (3,146 to be precise) were Japanese, 837 German, 222 Dutch, and 196 French. Furthermore, 90 percent of those that were more than one hundred years old had fewer than three hundred employees.” It goes on to give the examples of a German shoe store that for four centuries has always had only one location, and a Japanese hotel that “has been in the same family for fifty-two generations and even in its modern incarnation has only thirty-seven rooms”. I wish West had said something about banks, because some of them have been around a long time but, alas, he did not. Still, there’s something fascinating going on with these tiny businesses that know exactly what they want to do and never tempt fate by trying to expand.

## VII

Other thoughts: In order to make a reasonably-sized review, I’ve left out several sections. I didn’t talk about the Zipf and Pareto distributions that show up in a couple places. I glossed over a long summary of mid-century qualitative theorizing about cities from figures like Jane Jacobs, glossed over the discussion of mortality, and ignored altogether an account of Isambard Kingdom Brunel. Nor have I recounted the findings about how many different types of businesses exist in a city of a given size. There is a charm to Scale which comes from Geoffrey West’s meandering connections to a multitude of fields and persons, as well as the reach of his research. There is also risk inherent to having the summary of such an ambitious and boundary-less project being

authored by one man and placed in 450 pages. How can we be sure that everything discussed was presented fairly and appropriately fact-checked? I push back against West's opinions in some places, but at no point did I push back against his facts or get a sense of Gell-Mann effect. Nonetheless, the subject is broad enough that I can imagine some detail in here being wrong. This is a terrible criticism because I like the broad reach of the book, and given the broad reach there's absolutely nothing that could have been done differently so as to assure me on this front, but that's the world we live in. I'm willing to overlook it anyway, on the grounds that most of what's in the book is exposition of areas West has co-authored papers contributing to, and it's pretty clear which parts are flavorful or an aside.

I can imagine *Scale* receiving criticism from certain quarters because much of it is summary of existing knowledge rather than summary of original research. This doesn't bother me at all. I never would have heard of allometry or of the average number of heartbeats being constant across all mammals if not for this book popularizing the phenomena. Moreover, the format of the book places the contributions from West and his co-researchers into a narrative context where the reader understands what came before and what the thought process was as West moved from work in physics to work in biology to work in sociology. West is a strong believer in polymaths and multi-disciplinary work, and the structure of the book does a good job helping him make this case. And to comment here, I don't want to wade too deep into a debate on polymaths, but I know the word gets a bad rap in some places. I'll just remind those people that Geoffrey West isn't being a hobbyist here; he works for the Santa Fe Institute, and the research summarized here was published in various journals. If you want to know more, get ahold of a copy of the book and rummage through the appendix. Or read about the Santa Fe Institute; multi-disciplinary research is their shtick. While I'm on this subject, I'd also like to stress that West expresses plenty of appreciation and respect for the people who worked with him. He's not taking full credit for something that was a collaboration. I do get the sense that he took initiative in organizing the collaborations, but other than that he just happens to be the member of the team who wrote a book for the general public.

The one notable complaint I do have involves the discussions of sustainability. If one looks at the past couple centuries there is an obvious population explosion. Obviously, we haven't all collectively run out of food yet; we keep innovating and coming up with ways to feed the growing quantity of humans on the planet. But exponential growth is fast and terrible, meaning that we have to keep coming up with paradigm-shattering technologies. This provides evidence that a version of the Malthusian argument is going to be correct at some point. To this end West repeats the Kenneth Boulding quote that "Anyone who believes that exponential growth can go on forever in a finite world is either a madman or an economist." By itself I don't find this controversial. The assumption that there will be always be paradigm-shattering technologies to discover is

debatable, and then there is West's more technical argument that there has to be ever-decreasing amounts of time between those discoveries in order to prevent disaster.

The trouble I have with the Malthusian line is that the most prosperous countries aren't seeing explosive growth. The First World has birth rates that barely keep up with deaths. I'm open to arguments that this is not a permanent state of affairs, but Scale does not make any. Maybe this is because it is too busy responding to an Austrian School viewpoint of eternal optimism that I already disagreed with.

In the end this is a minor disagreement with the book and I'm glad that the book was written. I would love to see more literature along this line – math and physics style analysis of large scale human interaction. Certainly the results here are evidence that it's worthwhile to investigate.

# Scientific Freedom: The Elixir of Civilization by Donald W Braben

*(or maybe THIS is what Happened in 1971?)*

What if tech stagnation, declining growth rates, and the near-inevitable seeming collapse of the West are all because we got worried a few scientists would run off with our tax dollars?

That's the broad thesis behind *Scientific Freedom: The Elixir of Civilization*. Published originally in 2008, Scientific Freedom chronicles the journey of physicist Don Braben, as he designs and builds a Venture Research arm at British Petroleum in the 1980s. Braben was successful in funding a transformative research initiative at BP (transformative meaning it fundamentally changes humanity thinks about a subject). In his estimation, 14 out of the 26 groups funded made a transformative discovery, at the cost of only 30 million pounds over 10 years! A few examples of transformative discoveries made by groups funded by Braben in his time at BP are:

- Mike Bennett and Pat Heslop - Harrison discovered a new pathway for evolution and genetic control
- Terry Clark pioneered the study of macroscopic quantum objects
- Stan Clough and Tony Horsewill solved the quantum – classical transition problem by developing new relativity and quantum theories
- Steve Davies developed small artificial enzymes for efficient chiral selection
- Nigel Franks, Jean Louis Deneubourg, Simon Goss, and Chris Tofts quantified the rules describing distributed intelligence in animals
- Herbert Huppert and Steve Sparks pioneered the new field of geologic fluid mechanics
- Jeff Kimble pioneered squeezed states of light
- Graham Parkhouse derived a novel theory of engineering design relating performance to shapes and materials
- Alan Paton, Eunice Allen, and Anne Glover discovered a new symbiosis between plants and bacteria
- Martyn Poliakoff transformed green chemistry
- Colin Self demonstrated that antibodies in vivo can be activated by light
- Gene Stanley and Jos é Teixeira discovered a new liquid – liquid phase transition in water that accounts for many of water's anomalous properties
- Harry Swinney, Werner Horsthemke, Patrick DeKepper, Jean - Claude Roux, and Jacques Boissonade developed the first laboratory chemical reactors to yield

sustained spatial patterns — an essential precursor for the study of multidimensional chemistry

So how did Braben fund proposals, if he didn't use peer review or grant proposals?

## The Model

Don quite literally tried to “eliminate every selection rule imposed since about 1970 that appeared to stand in the way of freedom.” Don valued building a relationship, talking with the researchers, to determine whether or not they were of sufficient caliber to make a transformative discovery. Don and his small team were the end all be all. If Don got to know you, was impressed by your work, and thought you were working on something that was challenging and transformative, you got funded. Braben's conviction was what mattered, and he got results. He understood how difficult it would be for some of these folks to get funding (because of peer review). Don minimized overhead and administration by having minimal staff. For advertising, he would travel from university to university giving talks. He aimed to get to know the researchers at a personal level so that trust and rapport could be built in a way you can't do with a large funding agency. There was very little structure-no deadlines really to speak of, no reports to generate, just science.

In Don's approach, he never told anyone “no”. If he thought someone was a quack (for instance, if I claimed that I had disproven super-string theory with only knowledge of calculus), he would kindly probe deeper. “Will, you say you've disproven super-string theory, can you tell me how.” Don wouldn't tell me to buzz off. For a bumbling or nonsensical answer, he'd just tell me to come back when I had more. This practice let Don filter out fakers, without filtering too aggressively for ideas that might be true, but not accepted by the wider scientific community.

On a practical level, Venture Researchers would be funded for 3 years at a time. Support could be renewed and often was. On renewal, the director of the research program simply asked themselves whether what they were wanting to do was still challenging. If it was, the program got renewed. In this paradigm, trust between funder and scientists was paramount.

## Who Don Was Looking for: “The Planck Club”

Don is not egalitarian in his approach. Transformative Research not a program for everyone. He aims to fund researchers who can make the kinds of discoveries that are

paradigm shifting for humanity. Perhaps only 100 or scientists in a generation can make the kinds of paradigm shifting impacts Don is interested in finding. These scientists are what Braben refers to as the “Planck Club,” or the group of elite scientists who make the most notable discoveries of a given century. Here Don describes the early 20th century Plank Club:

The twentieth century was strongly influenced by the work of a relatively small number of scientists. A short list might include Planck, Einstein, Rutherford, Dirac, Pauli, Schrödinger, Heisenberg, Fleming, Avery, Fermi, Perutz, Crick and Watson, Bardeen, Brattain and Shockley, Gabor, Townes, McClintock, Black, and Brenner (see Table 1). However, I give this list only to indicate something of the richness of twentieth - century science. I wrote it in a few minutes, and it obviously has many important omissions. Other scientists would doubtless have their own. If the criteria for inclusion were based on success in creating radically new sciences, or of stimulating new and generic technologies, a fuller list could easily run to a couple of hundred.

The biggest breakthroughs often take a long time, and come from people with interests in all kinds of weird areas. Planck, perhaps Don’s favorite example, took 20 years working on Thermodynamics, and would never have made it if his research had been put under the pressure modern researchers are put under. Scientific Freedom, letting people work on their ideas without constraint, is essential to producing the kinds of discoveries the Plank Club made. In essence, Braben is allergic to bureaucracy.

## The Cost of Venture Research

One of the frustrating things for Braben, is the relative cheapness of Venture Research. In fact, if you believe Braben, the world is leaving trillion-dollar coins on the sidewalk:

The likely costs can be estimated using back-of-the-envelope figures. Let us assume that there were about 300 transformative researchers — the extended membership of the Planck Club — during the twentieth century. Let us adopt a rule of two by which we increase or decrease cost estimates by a factor of two whichever is the most pessimistic. Allowing for inefficiencies, therefore, let us increase the target number of transformative researchers we must find to 600 — that is, six a year on average over the century. This is a global estimate, but for a TR initiative in a large country such as the United States, then, according to our rule of two, we assume that all the new members might have

to be found in that country since it is home to about half of all R & D. If we also assume that the searches will be about 50% efficient, which Venture Research experience indicates would be about right, it would mean that a US TR initiative should find some 12 transformative researchers a year. (For comparison, a maximum of nine scientists can win the Nobel Prizes each year). TR is the cheapest research there is, as it is heavy on intellectual requirements but relatively light on resources. For Venture Research in the late 1980s – early 1990s operating in Europe and the United States, the average cost per project was less than £100,000 a year, including all academic and industrial overheads. Costs have gone up since then, so for our present purposes we might double them to, say, £200,000 or \$400,000 a year per project on average.

Transformative researchers should be supported initially for 3 years. Our experience indicates that about half of them would require a second 3 year term; and half of those, a third term of support. Very few projects should run for more than, say, 9 years. Those leaving the TR scheme either would have succeeded and been transferred to other programs created for them — that is, their research would actually have been transformative — or, the scientists agree that they had probably failed in their Herculean quests. However, these average figures are quoted for guidance; there should in fact be no hard-and-fast rules on the length of support. Remember Planck!

and

This sum should also be the steady total thereafter. As we have chosen  $x$  to be 12, after 9 years, therefore, a TR research budget (i.e., excluding overheads such as the initiative's administrative costs) would be some \$25 million a year. If it turns out to be significantly more than that, the initiative would be tackling a different problem than TR. After the first 9 years, the TR initiative would have backed some 108 projects, of which according to our experience about 54 should eventually turn out to be transformative in some way.

A TR budget for a smaller country — say, the United Kingdom — should be about half that of the United States, or \$ 12.5 million per annum. The Venture Research budget in our final year of operation (1990) was some \$ 5 million, two thirds of which we spent in Britain. As we had been operating for 10 years, it is possible that we had identified most of the researchers in Britain looking for potentially transformative research support at that time.

That's only \$25 million a year in inflation adjusted cost for a small country like the UK. They probably spend more on staples!



## What Changed

Pre-1970's, research was much smaller than it is now, and it was the norm that scientists could work on their problem of choice, without too much bother or oversight from their overlords. No moloch could touch these angels of knowledge, their tendrils of curiosity reaching out over nature, unencumbered by peer review.

Before 1970 or so, tenured academics with an individual turn of mind could usually dig out modest sources of funding to tackle any problem that interested them without first having to commit themselves in writing. Afterward, unconditional sources of funds would become increasingly difficult to find. Today, they are virtually nonexistent.

For Don, this change precipitated a decline in our ability to create breakthrough research. Peer review snuffed out all the weird people following their interesting passions. Instead of cool wacky scientists, we got salesman-scientists in suits. As people who can't get funding say, "*That dog won't hunt.*" There is a possibility that this change has precipitated our relative stagnation. Bureaucracy, and a lack of scientific freedom, the ability to get a small amount of unconditional money to follow your research interests, means that we don't get a Planck Club for the later half of the 20th century. Technology is the child of science, and if science is sick, maybe it makes us worse at creating the kinds of technology that keeps our world progressing towards a brighter future. Braben believes that although we have gotten many advancements in recent memory (the book was originally published in 2008), most of these are technologies leftover from the harvest of the early 20th century of research. This is important (and I think scores points for Don) because this gives him the title of being one of the earlier "alarm bells" of secular stagnation/decadence/tech stagnation in our society. Here, Don talks about the gift of the discoveries of the Planck Club:

"This prodigious progress came from our growing ability to harvest the fruits of humanity's intellectual prowess — scientific endeavour, as it is usually called. Material wealth continued to accelerate through most of the last century despite financial crashes and global wars. But then gradually, around about 1970, signs of major change began to emerge. Science's very success had unsurprisingly led to a steady expansion in scientists' numbers. That could not continue indefinitely, of course, and the inevitable crunch came when there were more than could adequately be funded. This was not only a numbers problem — the unit costs of research were also increasing. The funding agencies should have seen this coming, but they did not. Indeed, as I shall explain, many today do not accept

this version of events and are thereby contributing to one of the greatest tragedies of modernity. This perhaps surprising statement arises because the agencies' virtually universal response to the crisis was to restrict the types of research they would fund. Thus, to use a truly horrible word, they would prioritize, and focus funding on the most attractive objectives — that is, objectives the agencies perceived to be the most attractive. Thus, for the first time since the Renaissance, the limits of thinking began to be systematically curtailed.”

And

Thanks to that precious gift, and despite the havoc of world wars, financial crashes, and a threefold rise in population, per capita economic growth soared in the twentieth century, reaching a peak, coincidentally perhaps, around about 1970. It then began a steady decline.

## Conclusion

Why hasn't Venture Research caught on? I can only speculate, but I think letting folks run wild is *not* something that scales. Don might respond that it's perfectly okay that it doesn't scale-venture research is not for everyone, it's just for the select members of folks who have the capability to make transformative discoveries like the ones that belong in “The Planck Club.” It is important however, that *someone* is doing this kind of science funding.

We've many more researchers now than in the past, and there are simple bureaucratic reasons why oversight has become more important than research results. It's like building a vaccine-if you are a regulator, you don't get points for the hundred of thousands of lives you save, you only get punished if [365 folks get guillain-barré from your vaccine](#). The first researcher who gets public money, and spends “a little too much time down in Aruba” makes the front page of the Times, and the whole funding program is toast. On the bright side, it truly doesn't take much money to set up a venture research unit, and it's something that a rich tech founder could easily fund (Patrick Collison, are you still with me here?).

Scientific Freedom, for Braben, is something akin to the air we breathe. It's essential, but less obvious than water, health and security. It's tough to notice how important it is when you have it, but you sure as hell start to notice when it's gone. With it, society prospers, and we continue to find our own century's “Plank Club,” without it, we stagnate.



# Second Nature: A Gardener's Education by Michael Pollan

*On the use of land: A second look at Second Nature*

At some point in college I came by a copy of Michael Pollan's first book, *Second Nature*, about gardens and humans' relationship to the natural world. I may have stolen it from my parents' coffee table, or found it lying around the cozy, fervent MIT cooperative where I lived. Having been raised by former NOLS instructors on tofu, Thoreau, and Edward Abbey, and on the other hand undertaking at the time an intense education in technology, quantitation, and innovation (electronic paper, 3D printing, underwater drones, implanted medical devices...), the book stuck with me ('planted a seed', you might say), and played a part in leading me to this post.

Briefly, '*Second Nature*' presents contemporary environmental thinking about land as a sort of absolutist madonna/whore dichotomy between a few remaining preserved gems of 'pure' wilderness and the surrounding matrix of ordinary, 'degraded' territory, and Pollan pronounces this absolutist distinction barren – both literally and conceptually. As an alternative he sets forth the garden as a more fruitful metaphor, both for feeding ourselves and for thinking about interactions between humans and the natural world. Re-reading after over 20 (!) years, the book is thoughtful, engaging, and still has a good argument to make, one that resonates with my life.

Certainly Wilderness was a guiding star of my youth. Though my parents were raised in or near cities on both coasts, they met high in the Rockies as mountaineering instructors, and mixed in among the ordinary homesteading tools there were strange artifacts about – wood-handled ice axes, oval carabiners, and hanks of rattle-stiff old goldline. Practically before we could walk, my sister and I had full-sized Kelty framepacks waiting for us, and I remember the pride of finally being (marginally) big enough to carry mine. Many summers we road-tripped to Wyoming to hike high into the mountains,

disappearing across the continental divide where sometimes we'd go a week without seeing another party.

And since my parents 'settled down' pretty far off the beaten path, my sister and I grew up two miles from the nearest other kids, as part of [the broader back-to-the-land movement](#) which was in conscious opposition to the industrial practices of modern life. Growing up this way I came to see wilderness as Real, bracing and constant against the artificiality of television, plastic toys, and social cliques. The mountains we trekked through offered arresting beauty in reward for skill and hard work. The rock is enduring, the weather uncaring, the alpine trees strong and patient. Wilderness offers a tough but objective test – you either keep your gear dry or you don't, you make it over the pass or you don't, and the consequences flow directly from the nature of unvarnished reality. In wilderness it's clear that the universe doesn't care about you, but it's fair and its rules are legible – in that way it's far superior to junior high.

But, you can't cultivate a rock. You can't eat a view. As a species we are 7 billion people, ten thousand years down a one-way experiment in intensive food-making, culture-building, and technology-refining, and we've been [pretty darned sophisticated at it for thousands of years](#). For who we are, wilderness is an education, perhaps a vacation, but not a career – while it's certainly a nice place to visit, we just can't live there anymore.

Where we live is the industrial economy, and how we live there is mechanized agriculture and massive flows of energy. The surface experience of modern life in the developed world often obscures this – our skills are transferrable, our communities virtual, our finances digital. But this is, if not an illusion, an epiphenomenon of the stability, specialization, and efficiency of the underlying physical systems. Until we upload our consciousness into silicon (as the Singularity squad devoutly wishes), we remain stubbornly physical creatures – to see this clearly look no further than last spring's run on toilet paper.

The toilet paper thing is a sort of nervous joke, but I mean this in a broad and serious way. We don't actually have a post-industrial society; we just got so efficient at manufacturing buildings and cars and appliances and electronics that as a society we have enough bandwidth for many of us to toodle around with apps and stuff. Nor in physical terms do we actually have a post-agricultural society – we just got so efficient at agriculture (efficient in economic terms, by using a ton of fossil energy) that we had enough extra wealth/calories to build an industrial society. The foundation of the information age is hyper-efficient manufacturing, and the foundation of industry is hyper-efficient agriculture, and [the foundation of all of it is cheap energy](#). And the fact that the environmental impact of our lives is largely hidden by offshore manufacturing and high-voltage transmission lines does not make it go away.

Grappling in a real and quantitative way with human environmental impact seems to be what Second Nature is missing. Published two years before Pollan's book, Bill McKibben's *The End of Nature* introduced the US mainstream to concrete and present reality of global warming. But Pollan doesn't engage materially with climate change, or really with other large-scale environmental problems. Perhaps this is because he leans so heavily on the crisp dichotomy between wild and impacted land, which was at the very moment of his writing being ruptured by the all-permeating reach of global greenhouse gas emissions.

Pollan was right that wilderness is not a solution for sustaining 7 billion people, and that necessarily the way forward is to thoughtfully cultivate our world (enough of it to live on). But his book has precious little to say about actually producing sustenance. After a promising beginning among the truck farms of Long Island, the book takes long excursions into the social class implications of rose varieties, the moralistic overtones of compost, and the excesses of seed catalogues, and in the end seems to be more about aesthetics than substance, more about landscaping than actually producing

food, fuel, or fiber. His primary concern seems to be how a well-read suburbanite can display his good taste.

This disappoints me because I believe that the actual physical substance of how we live matters. To be sure, what we feel and proclaim matter also – aesthetics, symbols, and statements move minds, but minds are also subject to myopia, hypocrisy and wishful thinking. Rock-bottom physical reality matters at least as much, because all the while as we think and symbolize and post and upvote, inexorably we eat, we heat, we travel (or used to pre-pandemic), we buy, and we build, and the effects are real, quantitative, physical. Our microprocessors require electricity, our fingers quit typing much below room temperature, and our lofty professional and aesthetic goals are stubbornly dependent on a pound dryweight of bread, butter, and beans, daily with scant interruption. We remain tied, physically and therefore ethically, to the land.

Where then should we live, and how should we live there? Where? There's no point getting prescriptive about it; we are 7.6 billion now, we take up a lot of space already, and this Covid time is no season to encourage folks moving around. We should live in our communities. Urbanites have their own clear paths to low-impact living: density, bicycles, and a thoughtful diet go a long way. For those of us who live in the countryside, the risks and possibilities are broader.

I live in Maine. Because our state has the lowest population density east of the Mississippi, many of us live on sizeable chunks of land. While most people in the developed world externalize the environmental impact of their lives, still land and sunlight are the ultimate sources of our sustenance, and the impact does not go away just because we can't see it. So for those of us fortunate enough to own acreage, it's worth thinking about how our land could sustainably produce some of the basic stuff of human life.

I have written much here of my frustration with [the consumer model of citizenship](#). Not that I want to live as a survivalist, guarding a field of turnips

with an assault rifle, but nor am I satisfied to be merely a specialized cog in the global industrial machine. Because my life impacts the planet, because land and nature are the ultimate source of our sustenance, and because I have land, I am interested in stewardship. Because the path humanity is on is not sustainable, I am interested in experimenting and modeling other paths – ways to be productive, physically, of vital goods in a sustainable way and at a meaningful scale.

I say 'experimenting' because generally this sort of project won't make a whole lot of conventional financial sense; it's more like a hobby with a larger purpose. Why? Food is cheap, real estate is expensive, and the cash economy is lucrative. When a small buildable lot of an acre or so sells for \$50-\$100k, basically nothing (except cannabis) can be done agriculturally to match the economics of development. Still there is a sadness in old farms going to forest [or to subdivisions](#). To be sure, when the railroads spread across the nation, it made economic sense to move production of grain and beans from rocky New England farms to Ohio, Illinois, and Kansas, but nothing came to replace [the vitality that went out of the places then](#). But should the gods of the market dictate next that the staff of life should move on further, to be produced entirely in Brazil or Mongolia, will we then clothe Iowa in condos, and keenly await the grain ships [as the Romans did](#)?

In this modern first-world life, our food, shelter, warmth, transportation, and electricity all come from the global economy, which is 80% fossil-powered. Our land lies fallow as we heat our homes with petroleum, eat supermarket food grown with Haber-process nitrogen, and build with lumber trucked in from Canada or beyond. Those of us with the freedom to choose should contemplate instead how we might do better by thoughtful use of our land. What practices can I take up, such that if my neighbors and my bioregion followed suit, the result would be a stronger community, a more vital countryside, and a gentler impact on the broader world? What sustenance and beauty could we bring forth, and what might that do for our health, our communities, and our planet?



---

One interesting consequence of the intersection of ubiquitous internet search and ubiquitous aerial imagery is that it isn't hard to go looking and find places that would previously have remained literary abstractions. So without much effort I was able to find the western Connecticut property that was the centerpiece of Pollan's book. And in 2021 only the barest hints of the hardscrabble dairy farm he describes are visible. From the air and the street we have what appears to be a typical high-end exurban home in the woods, with more-tasteful-than-average landscaping. The only hint of anything out of the ordinary is a few raised beds, heavily shaded by large trees. Like so many former New England growers, Pollan has picked up and moved west, and the forest has largely reclaimed his efforts.

# Shattered Sword: The Untold Story of the Battle of Midway by Jonathan Parshall and Anthony Tully

## I: What is this book, and why am I reviewing it?

It's a book about the Pacific War of WW2, with a particular focus on the Battle of Midway. It's emotional, but highly detailed and technical, aimed at people who are already knowledgeable on the subject, or at least who have already read other books about it.

That is not me. I've really only read this one book on the Pacific war, plus various blog posts/wiki articles. And, uh, video games. I realize this is an area that is endlessly studied by both professionals and amateurs of all sorts, so compared to them I'm missing a lot. I expect a lot of push-back and technical criticism on this review, because I'm almost certainly getting some of the details wrong. Still, I do think that I at least know more than the average layperson, and I hope you can learn something from it too. Take it as an interesting story, both about the war and the vicissitudes of life in general.

I read this book because so many WW2 internet discussions recommended it. I kept seeing a pattern of like:

Person A: Long, detailed, plausible-sounding argument about the Pacific war, particularly about how the Japanese could have won at Midway

Person B: Have you read Shattered Sword? It totally debunks all your arguments

Actually, I've seen that a lot with WW2 internet arguments in general. I always found it frustrating. First, who has time to go read a full book just so they can respond to a short internet discussion? You're telling me to go read hundreds of pages before I can respond to a couple of paragraphs. Also, you have to go out and find some dusty old dinosaur book, when digital knowledge is flowing all around you on the information superhighway. Still, I kept seeing this *particular* book being recommended, and I really felt like all the free internet sources were lacking, so I gave it a shot and read a book.

(cue the Reading Rainbow theme song...)

Long story short, it's a good book. It's very technical and informative, but still easy to read for a layman. It combines both a high-level strategic view of the war with very human stories about the individuals involved. It lays out in detail the *quirkiness* of the naval forces involved (they're *kind of* modern, using steel ships and oil engines almost as fast as the ones today, but also not-manual targeting, no radar, and the "carriers" were often just wooden decks built on top of an old battlecruiser). It's an inspiring war story of a *huge* victory for the Allies, while still

acknowledging the suffering of the soldiers on the "other" side. It's clear that the authors were very diligent and did a lot of research- in particular, on Japanese primary sources that were previously unknown to American writers.

However, I found myself disagreeing with many of their conclusions, especially the ones that weren't quite stated directly. They start with a list of what they call myths about the battle, of which "correcting these distortions is the overriding goal of this book". They present themselves as being smart, dispassionate observers, correcting the mistakes of people who are caught up in emotions (Miracle at Midway is an example of the kind of book they're arguing against). In some cases there were also outright lies from the Japanese officers, which were used as primary sources by naive Western authors. Basically, they argue that the results of the battle were not a "miracle" at all, but rather to be expected, and almost inevitable when properly understood.

Oddly, the more they argue against these "myths" the more I found myself *believing* them. At least the essence of them, if not the specific details. You could call it a form of the "Streisand Effect". The authors assume that everyone reading the book has bought into those "myths" based on Hollywood movies and flawed earlier works. Personally, I didn't even know the myths, except insofar as they showed up in Star Wars movies (more on this later), so their "corrections" actually introduced the myths to me, and seemed more like nitpicks than takedowns.

Here's the full list of what they call myths:

- The Americans triumphed against overwhelming odds at the Battle of Midway.
- The Aleutians Operation was conceived by Admiral Yamamoto, the commander in chief of Combined Fleet, as a diversion designed to lure the American fleet out of Pearl Harbor.
- During the transit to Midway, Admiral Yamamoto withheld important intelligence information from Admiral Nagumo, the operational commander of the carrier striking force. As a result, Nagumo was in the dark concerning the nature of the threat facing him.
- Had the Japanese implemented a two-phase reconnaissance search on the morning of 4 June, they would have succeeded in locating the American fleet in time to win the battle.
- The late launch of cruiser *Tone*'s No. 4 scout plane doomed Admiral Nagumo to defeat in the battle.
- Had Admiral Nagumo not decided to rearm his aircraft with land-attack weapons, he would have been in a position to attack the Americans as soon as they were discovered.

- The sacrifice of USS *Hornet*'s Torpedo Squadron Eight was not in vain, since it pulled the Japanese combat air patrol fighters down to sea level, thereby allowing the American dive-bombers to attack at 1020.
- Japan's elite carrier aviators were all but wiped out during the battle.

I'm not going to go into detail on all of these like the book does, but suffice to say they all seem more like "exaggerations" than actual myths, even based solely on the arguments presented in the book. I was left with the distinct impression that, yes, the battle *was* won against the odds, and it was only through a combination of astoundingly bad Japanese command decisions and heroic sacrifices by American pilots that it was won at all, let alone as thoroughly and decisively as it was.

## II: The gritty details

The bulk of this book is essentially a play-by-play of the battle itself, with commentary mixed in. It's very detailed, very thorough, and hard to summarize- this is what the bulk of the book is about. I'm going to have to skip over a lot of this for the sake of brevity. Here are some representative quotes:

Back on board *Kido Butai*, the possibility of imminent air action had spurred *Soryu* into launching three fighters to augment the combat air patrol (CAP) at 0600. *Hiryu* followed suit at 0612.

...

Despite the hell they had just come through, Fleming's men managed to bracket *Hiryu* with numerous near misses between 0808 and 0812, some as close as fifty meters from the ship.

...

At 0917, with the strike aircraft being struck below, Admiral Nagumo altered his course slightly to 070, still at battle speed 3 (twenty-two knots), to close the enemy, which he believed lay directly ahead.

It's overflowing with details about individual ships, aircraft and personnel. I'll try to summarize it in a way that just gives the main points, but I wanted to give a taste of how much detail went into this book. I'm going to have to gloss over most of that detail for the purposes of this review.

Should I give a spoiler warning? If you didn't know, the battle of Midway was a HUGE victory for the Americans, and for the Allies in general. From Pearl Harbor(December 7, 1941) until Midway (June 4, 1942) the IJN was "in the lead", having both more and better ships than

anything the US had available in the Pacific. It ran a series of long range offensive missions, almost all of them successful. It sank ships from the British, Dutch, and US navies, while invading most of Southeast Asia. The Battle of Midway completely reversed that position- after that battle the US largely controlled the open ocean, while Japan could only defend near islands it controlled or launch night raids.

Midway itself is a very small island, mostly underwater, with just enough land for a single airstrip. It's only strategic value was its position: 1,800 km west of Hawaii, roughly halfway between the US and Japan. The Pacific ocean there is remarkably empty, so there was nowhere else to put a military base.

The battle was mostly fought, and won, by aircraft carriers. Partly this was random chance: The US had all of its battleships sunk or damaged at Pearl Harbor, leaving only its aircraft carriers available (this was a lucky fluke, since they just happened to be away on a training mission when the attack came). But it was also a result of how dominant and effective carriers were at this time- the IJN had several battleships and other heavy surface ships available, but still relied mostly on its carriers. Both sides had submarines, but they were quite slow and limited. You can't *totally* ignore the other ships involved, but you can simplify a lot by focusing solely on the carriers.

This applies to analyzing the results of the battle, too. The Japanese brought 4 aircraft carriers to Midway, and lost all four. The Americans brought 3, and lost only one (which was actually sunk by a submarine after the main battle had ended). I'm going to borrow Scott's Simplicio/Sophisticus device to talk about these results, paraphrasing how the book describes it:

Simplicio: Splash 4 carriers! Boom!

Sophisticus: Don't be so simplistic. Sure, a carrier is a big obvious loss, but the more important losses are intangible. Morale, the loss of experienced pilots, the unit coherence of an experienced carrier air group, the loss of strategic momentum, etc...

Simplicio: No, don't overthink this. The really important loss was the carriers themselves, especially the big fleet carriers(as opposed to escort carriers, which were much smaller and more limited). The Japanese had very limited shipbuilding capacity compared to the allies. They built most of their ships with help from British ship designers and resource imports from before the war. During the war, they struggled to build ships, especially really big ships like fleet carriers. Fleet carriers(as opposed to small escort carriers) were the only ships capable of launching a really effective air strike. They built a grand total of *one* purpose-built fleet carrier during the war, the *Taihō*, which was started before Pearl Harbor and not finished until 1944. The Japanese had a lot of other pilots and other aircraft, and continuously produced more of each during the war- they just didn't have any good place to launch them from, especially when trying to fight over small pacific islands with no good runways. Without carriers, the Japanese basically couldn't bring sufficient aircraft to any fight that wasn't near their home islands.

It's strange to think of the US navy ever being behind in ships. It massively outproduced Japan (and every other navy *combined*) during the war. But in 1942, none of those ships were *finished* yet. All it had actually available were relatively small numbers of ships built before the war, some of them quite old or with bad designs. The Japanese, on the other hand, had gone all out to build up their navy in preparation for the war, and had attacked Pearl Harbor at the peak of naval readiness. Here's how the book describes the situation:

In this modern era of unchallenged American naval supremacy, merely restoring parity may not seem like much of an accomplishment. But it must be recalled that Midway was fought "between one navy at the peak of its strength and another if not at its nadir then close to it."<sup>23</sup> In the dark months of 1942, being able to claw back to parity was an enormous achievement.

In terms of fleet carriers, the Japanese actually had a *lead* in raw numbers at Midway: four for them, vs three for the US. You could count the island of Midway itself as a carrier too (it's a very small island, practically a coral reef, with just enough room for a single airstrip and roughly the same number of planes as a carrier). But that still just puts the numbers even-stein. It almost seems like fiction, or like something from a computer game: "4v4@Midway, no noobs".

It's something of a stroke of luck for the Americans that they even managed to have *even* numbers at Midway. The authors go into quite a bit of detail on this subject, arguing that it was a mix of incompetence and arrogance on the part of the Japanese navy. They had started the war with a total of 8 large carriers. 6 of them were used for the attack on Pearl Harbor. They assigned two carriers on a totally pointless mission to attack one of the US Aleutian islands, which is now known as the "Forgotten Battle". Two others were damaged a few months before Midway, at the Battle of Coral Sea- they were eventually repaired, but the Japanese showed no urgency in getting them ready for Midway at all. This left only four- just half!- of their large carriers for this battle, despite it being a battle of critical importance. They also had several smaller carriers which were not used.

To put this in perspective: let's imagine that the US was faced with a HUGE emergency. Like some sort of pandemic, threatening to rapidly infect the entire nation and kill massive numbers of people. And there was a way, albeit difficult and dangerous, to rapidly develop a vaccine or distribute it faster but the bureaucracy in charge totally denied it, instead insisting on following all the standard rules and regulations to the letter, unwilling to try anything at all unconventional to speed up the process. Good thing that would never happen here!

Meanwhile, the USS *Yorktown* (also damaged at Coral Sea) was estimated to need 90 days of repair: The US Navy did it in three, by working nonstop and cutting a lot of corners, because they understood that this was an absolute, no-holds-barred emergency. To make a long story short: despite its massive advantage in shipbuilding and industrial production, it was rather fortunate that the US could even manage *even* numbers of carriers/runways at Midway (4 each), rather than being significantly outnumbered.

And of course, raw numbers don't tell the whole story. Quality matters just as much, if not more. The IJN flight crews were all *extremely* experienced. Japan had been at war (formally) with

China since 1937, and informally even before that. Their carrier crews trained extensively for the attack on Pearl Harbor, and had followed it up with attacks on British, Dutch, and American targets all around the Pacific, almost all of them successful. The US had been unexpectedly pulled into the war with the surprise attack on Pearl Harbor, and had only one significant battle since then. Coral Sea). Compared to the IJN, the US Navy was a bunch of rookies at this point, especially in carrier strike operations. Here's how the book describes the difference:

Japan's carrier force in particular was truly without peer. At Pearl Harbor it demonstrated a level of sophistication that the U.S. Navy would not be able to replicate for another two years. Whereas the Allies were still using their flight decks singly or in pairs, Japan had used six fleet carriers to sweep American airpower aside and smash a major naval base in broad daylight. In terms of their ability to use massed airpower, the Imperial Navy had no rival. Japan's pilots were war hardened, supremely aggressive, and highly skilled. Likewise, Japanese carrier aircraft—epitomized by the marvelous Mitsubishi Zero—were in many cases superior to those used by the U.S. Navy at this stage of the war.

The IJN had an advantage in weaponry, too. At the time, there were essentially three ways of sinking a ship.

- Really big guns, like the kind on a battleship
- Torpedoes, fired either from a submarine or from a torpedo-bomber plane
- Bombs, dropped either high above from a level bomber, or from a dive bomber swooping down close to the target

The IJN had proven capable at using all three of these. There was Pearl Harbor, of course, when it used dive bombers with deadly accuracy. It used torpedo bombers to destroy the only British battleship and battlecruiser in the Pacific just three days after Pearl Harbor, and then a submarine torpedoed a US carrier, the *Saratoga*, in January, which prevented it from taking part in Midway. Its battleships didn't have much to do, but they were still highly capable and feared.

The USN, on the other hand, was not. First, all of its Pacific battleships were sunk or heavily damaged at Pearl Harbor, so that option was completely gone. Its torpedoes were the victim of the Great Torpedo Scandal, which led to some alarming incidents, like hitting targets and not exploding, or even turning around to fire on their own ship (!):

"It tended to run "circular", failing to straighten its run once set on its prescribed gyro-angle setting, and instead, to run in a large circle, thus returning to strike the firing ship. "

One US submarine, the *Nautilus*, actually caught sight of the Japanese fleet before Midway, and had a chance to attack a battleship. But it suffered some difficulties:

Brockman [the captain], though, was having problems setting up his attack. One of his torpedoes was running hot in its tube, having had its retaining pin sheered away during the depth charging. It was making a hellacious racket, and Brockman was

certain that the Japanese escorts could hear its banshee wailing. Nagumo's fleet—what little Brockman could see of it—was still on a westerly heading when *Nautilus* fired at 0825. His target was *Kirishima*. Taking aim at her starboard side, Brockman let fly with two torpedoes at a range of 4,500 yards. Or at least he thought he did—he found out later that one tube did not fire, leaving only one fish streaking toward the target.

The US navy's torpedo bomber was, to put it mildly, flawed:

The Devastator proved to be a death trap for its crews: slow and hardly maneuverable, with poor armor for the era; its speed on a glide-bombing approach was a mere 200 mph (320 km/h), making it easy prey for fighters and defensive guns alike. The aerial torpedo could not even be released at speeds above 115 mph (185 km/h).[19] Torpedo delivery requires a long, straight-line attack run, making the aircraft vulnerable, and the slow speed of the aircraft made them easy targets for the Mitsubishi A6M Zeros.[20] Only four TBDs made it back to Enterprise, none to Hornet and two to Yorktown, without scoring a torpedo hit.

In terms of actual, functioning weapon systems available at Midway, the USN only really had bombs. And here again, there were issues. Dropping bombs from a high altitude (like from a regular bomber) was too inaccurate and slow- the IJN carriers could simply watch them falling and steer somewhere else, as seen in this photo. You don't normally think of a huge ship like an aircraft carrier "dodging", but compared to a primitive bomb being dropped from 10,000 feet, it could dodge very easily.

That leaves dive bombing, a hair-raising technique of flying almost vertically down, only to pull out at the last minute and release its bomb directly on top of the target. Here's how the Germans tested their dive bombers:

It was discovered that the highest load a pilot could endure was 8.5 g for three seconds, when the aircraft was pushed to its limit by the centrifugal forces. At less than 4 g, no visual problems or loss of consciousness were experienced.[29] Above 6 g, 50% of pilots suffered visual problems, or *greyout*. With 40%, vision vanished altogether from 7.5 g upwards and black-out sometimes occurred.[30] Despite this blindness, the pilot could maintain consciousness and was capable of "bodily reactions". After more than three seconds, half the subjects passed out. The pilot would regain consciousness two or three seconds after the centrifugal forces had dropped below 3 g and had lasted no longer than three seconds. In a crouched position, pilots could withstand 7.5 g and were able to remain functional for a short duration. In this position, Junkers concluded that 2/3 of pilots could withstand 8 g and perhaps 9 g for three to five seconds without vision defects which, under war conditions, was acceptable.[31]

This sounds extremely difficult and dangerous! It also sounds like something you'd need a lot of training and experience to handle well. Even before the war, it was seen as difficult and dangerous, worthy of a Hollywood movie starring Errol Flynn. Unfortunately, the USN had only gotten its dive bombers in 1941, giving little time for their pilots to learn.



There were also fighters, designed to shoot down enemy planes rather than attacking surface targets. The Japanese had the Zero: fast, long range, and highly maneuverable, although lightly armored. The Americans had the Wildcat: more armor, but slower and less maneuverable. A decent plane, but here's how the book describes the balance before Midway:

To this point in the war, aerial combat had been pretty much a one-way street for Zero pilots—they dished it out, and the enemy died.

Suffice it to say, the Japanese planes at Midway had a qualitative advantage in almost every respect. The authors of Shattered Sword argue that it was a myth that the Japanese had overwhelming numbers, because only a small portion of those numbers were actually used at Midway, and so the Americans had equal or greater numbers of aircraft at the battle. That's true, but it's worth remembering how poor the American aircraft at the battle were. Out of about American 350 aircraft, the only ones actually capable of sinking a Japanese ship were the 100 or so carrier-born dive bombers with mostly limited experience. The IJN had only 250 aircraft, but all of them were proven to be highly capable machines, while their crews also had more experience and training. That might not be "overwhelming", but it was a large advantage.

The battle itself happened in stages, as various groups of aircraft encountered their targets. It started with the Japanese bombing the island of Midway itself, heavily damaging it and destroying most of its marine corps fighter defenders. Shortly after, the bombers from the island attempted to attack the Japanese carriers, but couldn't do much- they either missed or were shot down. After that, the American carrier-launched planes arrived. First, the torpedo bombers were all shot down, while the American fighters were too slow to offer protection. Finally, the dive bombers arrived, which managed to break through and deal critical damage to three of the four Japanese carriers. The surviving Japanese carrier launched a counterattack which heavily damaged (but didn't quite sink) one American carrier, and was itself sunk by another American carrier strike later that day, again delivered by dive bombers.

I'm skimming over a *lot* in that one paragraph- that's about half of the book! The authors argue at length about why individual commanders acted as they did, in particular the Japanese captains and admirals. There's a lot of discussion about, say, why wasn't the Japanese air defense better, or why didn't they launch their own strike sooner. They do a good job of justifying all of the individual decisions, making it seem like they were all either lacking key information, or following protocol, or just didn't have time to think properly. For example, this is how they explain the disorganized Japanese air defense:

Nagumo was stuck on board a wildly maneuvering carrier, watching his other vessels running pell-mell in all directions. Every time it looked as if things were settling down a bit, another air raid warning would come in. Nagumo can hardly have known where all his *ships* were at any given time, let alone have had an appreciation of what his aggregate CAP strength was.

The implication from all of this commentary on command is that victory was mostly inevitable for the Americans. The Japanese did everything about as well as they could, but there was just no

stopping it. And they do argue this very well, with a ton of details and primary sources that I'm not at all qualified to debunk.

But that won't stop me from trying! Let's go back to that battle summary. To reiterate, *all* of the Japanese carriers were sunk by American dive bombers- the level bombers simply missed, and the torpedo bombers were shot down like flies. One of carriers was destroyed entirely by *one guy*, Richard Best, who had ~~the awesome power of nominative determinism~~ far more experience, training, and skill than most American pilots (he had served on one of the very first US carriers, been a flight instructor for a while, and then specifically requested a transfer back to being a combat pilot again). There was a screw-up during the attack, which led to *all* of the American dive bombers targeting just *one* Japanese carrier- Best was the only who realized the mistake and went after the other (with his wingmen), and he was the only one who hit it, doing fatal damage with just the one bomb (it started a fire which burned out of control). The carrier he hit, the *Akagi*, was in the process of launching its own strike against the American carriers, which it almost certainly would have sunk, given how good the IJN was at that sort of thing (they were bad at defense, but the world leader at offensive air strikes). Later that day, Best attacked the remaining Japanese carrier, *Hiryū*, and there's at least one witness who says that he was one of the few that hit it.

In other words, this one guy, by himself, managed to

- definitely destroy one Japanese carrier, the *Akagi*
- *probably* save an American carrier, *Enterprise*, from destruction
- *possibly* lead to the destruction of another Japanese carrier, since it was the planes from the *Enterprise* that sank her, with Best among them

So a lot of the battle's key results simply come from the actions of this *one guy*. I still can't get over how incredible that is. You usually hear arguments about whether battles were decided by strategies and tactics from the commanders, or larger technical and economic factors. But here, it's just one individual soldier being really good at his job. As an "individual contributor" myself, I can't help but find that inspiring.

Also, I can't waste the chance to use this quote:

In the words of his backseater, Aviation Chief Radioman James F. Murray, "Nobody pushed his dive steeper or held it longer than Dick."

Sorry. I'll now return to the high-brow, sophisticated content that SSC readers expect.

### III: Analysis and Takeaways

The book itself kind of downplays Best's actions. They agree that it was him who scored the critical hit on the *Akagi*. But they spend more time arguing against myths in other descriptions of the battle. One of the myths that the book sets out to debunk is the "Five Fateful Minutes"- a story (written by Mitsuo Fushida, a Japanese pilot) that the *Akagi* was only five minutes away from launching its own strike, so Best's hit came at the perfect time when its deck was full of

fully-loaded planes about to take off. Apparently this was true, it would have taken "at least thirty" for their strike to be ready, and the planes were still below deck.

But I think it hardly matters- 5 minutes or an hour, planes on deck or below, that's still a *very* narrow window of time that decided the battle, all from *one* bomb hit. Fushida might have been wrong about the details, but I think he was right about the essence here. If Best had missed, or had simply followed orders instead of taking initiative for himself, there was no one else who could have destroyed *Akagi* before it launched its attack. And while it's well known that *Star Wars* borrows heavily from WW2 air battles, this sounds almost exactly like the plot of the movie! The Death Star is just about to fire its Superlaser that would destroy the Rebel base, and Luke has just one chance to make a precisely aimed hit to destroy it before it can fire. Points to *Star Wars* for being... historically accurate?

Another point of comparison to *Star Wars* must be made. In *Star Wars*, the enemy is some sort of evil empire ("The Empire"). Imperial Japan was a *real life* evil empire. I'm not going to be neutral here- The US did some bad things too (in particular, its internment camps and its bombing attacks on civilian targets later in the war), but nothing on the same level as what Imperial Japan and its military did. They were invading every country they could, massacring millions of civilians, and committing atrocities so horrible that I don't even want to think about them.

The tone of this book, on this issue, is surprisingly neutral. They mostly gloss over the war crimes of the Japanese military, make fun of other works which call the US "good guys" and do some amateur psychoanalyzing to try and justify what the IJN was doing. Here's one representative section:

To the average Westerner, steeped in the winner's history of World War II, any attempt to justify Japan's war in terms of Pan-Asian liberation is simply so much hogwash. The Japanese were aggressors, the Allies liberators, and everything from a moral standpoint has been very much cut and dried for half a century. The prevailing American attitude toward the war was crystallized as soon as the first Japanese bomb fell on Pearl Harbor. Yamamoto's "sneak" attack simply put an exclamation point on the writ of contemporary American moral outrage over previous Japanese aggressions in Asia.

Yet, despite the fundamental validity of these Western views, it is important to recall that at some level the Japanese people sincerely believed they were fighting for a larger cause, whose intrinsic good was undeniable. If they were also capable of ignoring the social injustices and outright atrocities—which were many and sordid—that accrued under this banner, then that sublimation came about from a conviction that achieving the larger goal of destroying Western colonialism somehow justified the means employed. This long-standing rationalization lies at the core of Japan's inability to examine and condemn its own wartime actions with anything approaching the sincerity and candor that its victims feel is required.

I *strongly* disagree with that view. Perhaps it had some merit in the past as a reaction against the movies and popular fiction of the 1950s which portrayed the American military as glorious heroes who could do no wrong. But it is simply incorrect to make it seem as though the aggression and atrocities of the Japanese military was in any way justified. They were *evil* (as a whole, not every single individual of course), and any victory against them deserves to be celebrated just as much as the storming of Normandy on D-Day or the Battle of Stalingrad. There should not be any moral hesitation in openly celebrating the American victory at Midway.

Going back to Science Fiction- you know that ship from *Star Trek*? The *Enterprise*? It was named after a real ship- the Enterprise! It's the most decorated ship in US history, and arguably the single most effective warship of all time. For a while, after Midway, the *Enterprise* was the *\*only\** functioning US carrier in the Pacific, which led to an iconic picture of its crew with a sign saying "Enterprise vs Japan." Quoting Wikipedia on its service record: "By the end of the war, her planes and guns had downed 911 enemy planes, sunk 71 ships, and damaged or destroyed 192 more." That ship utterly wrecked the IJN, and it sometimes did so with very little support. That's the ship that launched strikes that sunk two carriers and Midway, and likely would have been destroyed if the *Akagi* had been able to launch its strike. I find it interesting that the ship in *Star Trek* is portrayed as a peaceful, humanitarian vessel that only uses its weapons as a last resort. The real life *Enterprise* was the opposite of peaceful, but it still deserves to be celebrated.

Here's the part where we play what-if. What if the Japanese had won an overwhelming victory at Midway, sinking all the American carriers while losing none of their own? Could that have led to them winning the war?

No, not a chance. The book goes through this quickly at the end, and dismisses it as insane. ("At most, the defeat at Midway cost the Japanese approximately eighteen months of strategic leverage that their four carriers might have bought them.") However, it would have been a *longer* war, and *many* people would have died each single day that it went on. That's why it's important to celebrate the victories, without thinking too hard about weird alternative histories. The more likely alternative history is that the US could have avoided being aggressive, and simply waited for all their new ships to finish being built- it's heroic, and commendable, that they were as aggressive as they were despite their lack of numbers.

A more reasonable question would be, could the Japanese have won this particular battle? The authors make it seem like they couldn't. They never quite come out and say so directly, but whenever they discuss the decisions made by Japanese commanders, they make them seem reasonable and inevitable given the circumstances and their overall strategy, doctrine, and culture.

But then the book admits that, yes, all of those decisions ended up losing them the battle. It almost seems like every single decision they made regarding the battle was wrong:

- Why attack at all? Midway island itself wasn't valuable.
- Why not realize their codes had been compromised? Or have better codes to begin with.

- Why attack with *half* their carriers? Sending fewer (or none) would have been less of a loss. Sending more would have had overwhelming numbers in the air.
- Why send the carriers charging in first, while the battleships and other ships hung back too far to do anything? Or even send them in ahead?
- Why put up just a minimal amount of scouts? They didn't have radar, so they really needed more scouts. Instead they were totally caught by surprise.
- Why was their fighter defense so disorganized? They just sort of flew around randomly, and eventually got caught completely in the wrong position.
- Why were their ships so vulnerable, especially to fire?

There are reasons for all these given in the book, but they're not especially *good* reasons. They all basically amount to either "that's just the way things were done in Imperial Japan and its navy, and no individual commander could possibly change that" or "under the stress and chaos of the situation, no one could have known better." E.g., part of the reason they were so vulnerable to fire was that only officers were allowed to know how to fight fires, not regular crewman. They attacked with only part of their fleet, because they thought the Americans would be too scared to come out and fight if they sent everything. Seen in that light, where no decisions can be changed, I suppose it's true that the outcome was inevitable. But it seems odd that a victory which depends on the opponent screwing up everything in the worst way could be seen as "inevitable". And even despite all those bad decisions, it still ended up being decided by a small number of bomb hits by American dive bomber pilots, particularly the one by Best on the *Akagi*.

So I came away from this book conflicted. The authors clearly did a massive amount of research on the battle, and they managed to convey a ton of technical detail in a way that was still readable. They persuasively argued all of their individual points. But there was still this not-quite-outright-stated tone to it that I found myself disagreeing with. Perhaps some earlier books on Midway had gone a little too far in hyping it and they wanted to correct them, fine. But they went too far in the other direction to make it seem almost inevitable and not worth celebrating. Even with all their arguments to the contrary, Midway still seems like an incredible victory to me.

# Shut Out by Kevin Erdmann

## I.

I have long believed in a paradox of popular narratives. When multiple contradictory narratives emerge around a topic of import to broader society, they all contain parts of the truth (fringe conspiracy theories notwithstanding). But when a dominant society-wide narrative emerges it is usually fundamentally false. For instance, when crime was precipitously falling from the mid-90's to the mid-20-teens, poll after poll showed that up to 90% of (bipartisan) Americans perceived crime rates at least level over that period, with a majority seeing them as getting worse.

There was a whole genre of popular explanations for the housing “bubble” and Great Recession. Though they varied on the specifics, they all shared the same familiar premise: there was a speculative bubble in real estate fueled by “irresponsible” lending which resulted in over-building; when it popped, as all bubbles inevitably do, leveraged banks became insolvent and either failed or needed propping up; the aftermath led to one of the deepest recessions in living memory.

In these polarized times when we as a society cannot agree on very basic facts, one would be hard pressed to find anyone who lived through that time and would not be on board with the above description of the economic crisis (what should have been done about it is another story). But in his book [Shut Out](#), author Kevin Erdmann persuasively argues the dominant narrative that has emerged is incorrect, and rests on a series of myths that appear superficially true, but collapse under careful scrutiny.

His thesis is that the run-up in housing in the mid-2000's was not an asset bubble superimposed on the crest of a cycle, but a regime shift to higher housing prices caused by supply restrictions in the most economically prosperous and dynamic US cities, which he has deemed “closed access cities” (and include NYC, LA/SD, Silicon Valley, and Boston). “Loose” credit was a necessary facilitator of people moving into and out of these cities and their unique housing markets, was more effect than cause of increased housing prices, and did not focus on buyers with abnormally low credit rating or income. People left these dynamic areas preferentially for more affordable housing in rapidly growing areas in the Southwest and Florida, deemed “contagion cities” (including Las Vegas, Phoenix, inland California, and Tampa). And finally, the subsequent bust and recession were not foreordained consequences, but were imposed by monetary and

regulatory policy, with various levels of intention. These first triggered fear in the housing market ending the boom, then allowed a deep recession to proceed that set off a financial panic, and finally, resisted a recovery in the housing bust which had by then spread to the rest of the country and would last well into the next decade.

In short, there wasn't a demand bubble, but a supply problem; the banking shenanigans were a bit player; and the federal reserve misidentified the problem and responded with a series of disastrous decisions.

## II.

These are pretty audacious claims. Ben Barnanke is seen by many as a hero who competently steered us way from an even deeper crisis. *The Big Short* was a best-selling book made into an Oscar-winning film. Major legislation and new federal entities were created wholly on premises that were the antithesis of the conclusions of *Shut Out*. Occupy Wall Street, to my knowledge, never tried to Occupy the Zoning Board.

So, Erdmann has to really deliver the goods, and deliver he does. It is one of the more relentless data-driven exercises in upending conventional wisdom I've encountered. It simply notes that the credit-fueled bubble thesis necessarily implies certain facts, and then looks at local markets individually to see if the data support that story. And time and time again, they do not.

There are basic facts that would have to be true to support the conventional story: 1) the large price increases should be widespread throughout the US; 2) given that mortgage-credit markets are not international, the US run-up in prices would have to be unique; 3) homeownership and home-building should be unusually high; 4) mortgage securitization should have led or coincided with homeownership; 5) the quality of borrowers and mortgages (credit and income) should have decreased; 6) the initial bust should have hit the low-income owners the hardest; 7) foreclosures should have happened soon after the bust; 8) and given the regulations that arose from post-crisis legislation, the prices should not have risen again to those heights. None of these are accurate.

(1) The price increases that made headlines were a hyper-local phenomenon. Erdmann only looked at twenty metropolitan areas, but among them, only those in the "closed access cities" on the California and northeastern coasts and the "contagion cities" in the Southwest and Florida did prices even resemble a bubble. In the other metro areas he surveyed, and in the suburban and rural areas of the country he did not, price increases were modest or non-existent.

(2) The same price behavior was seen in all English-speaking countries in the early/mid 2000s, and all except the US and Ireland saw a quick recovery to that high price trend after the recession (Australia's was unabated because they even avoided the recession). To be sure, the US was unique, but it was in the price behavior after the bust, not in the boom, that stands out. US policy- and credit-centric stories fail to explain the housing markets in Australia and the UK because there wasn't really a "bubble".

(3) Homeownership, adjusted for age, was not increasing; new housing, adjusted for population, was being produced just above its long term average, nowhere near its high. These are often confused, because of composition effects. Homeownership rates did increase several percentage points, but this was a function of an aging population. Single unit-housing starts were at a peak, but this was partly due to a permanent shift away from multi-unit apartments and manufactured homes. Total housing units were being produced at a very modest rate given the increased prices.

(4) Mortgage securitization held by private institutions (which represent subprime loans) severely lagged the increase in homeownership, and first-time homebuyers as a percentage of all homebuyers actually decreased during the period from 2003-06 where these loan products took off.

(5) The marginal buyers coming into homeownership had high income (see more in section below), not low. At each income quintile, leverage was decreasing.

(6) During the initial bust, the rise in delinquencies was concentrated in the population with high income and credit scores; many were investors who were strategically defaulting.

(7) Vast majority of foreclosures happened well after the bust because they were due to the larger recession (and occurred in areas that did not see a run up).

(8) Prices have resumed their upward trend in the face of a prolonged credit bust in areas that had seen the biggest price increases during the boom, because it was fueled by supply restriction.

### III.



It would be impractical to relay all the data that support Erdmann's thesis, but I'll go into one example to illustrate. Again, the conventional story is that credit was expanded to underqualified homebuyers who could only justify a given mortgage because home prices were going up. Clearly, if this were true you would expect to see marginal buyers in places where home prices were increasing the most be lower-income. But this is not the case.

Net domestic migration in the "closed access" metro areas between 1995 and 2008 was a) negative, and b) the mirror image of that seen in the "contagion cities." At the height of the housing boom in 2004, these closed access cities lost 1.5% of their populations, and the contagion cities gained 1.5%; afterward, these trends reversed as housing prices rose more sharply in the contagion cities, and then prices everywhere fell during the recession.

This relationship between the migration statistics in the two groups is not a coincidence. (Shocking facts: if you were a resident of Boston or New York around 2000, and you moved to a newly built home by 2006, you were as likely to be living in Florida as to still be living in NY or Boston. More homes were built in AZ/NV/OR just for former residents of SF/LA than were built in SF/LA at all). And the timing of the changes in each category are a major hint that there was no speculative bubble. Erdmann writes, "Doesn't it seem strange that every location supposedly caught up in a speculative frenzy is characterized by thousands of residents moving away as home prices rose? If the problem in the Contagion sites was too many homes, why would these cities see a sudden drop in in-migration and an increase in out-migration [in 2006]?"

And this sudden shock to inter-city migration in 2005-6 explains the pictures you probably saw of empty homes and neighborhoods. For years, Phoenix had been issuing building permits at an unprecedented pace, but this barely kept up with demand from rising in-migration, *as prices still rose significantly*. The migration shock was sudden, akin to the music stopping in a game of musical chairs. But in this game, chairs were appropriately being manufactured because the number playing in the game had been increasing, and when that stopped (temporarily), there was a temporary overhang in these markets.

Census data demonstrates that the population of people that had been leaving the closed access cities (and entering the contagion cities) consisted primarily of two groups: homeowners with well-above average incomes who were cashing in their increased equity, and renters with below average incomes who were escaping to affordability (and were still renting, meaning these new units were investor-owned properties). The people moving into and buying in the closed access cities also had above average incomes (and high educational attainment).

None of these groups that were on the move in this period, either of the high income populations or the investor-landlords, fit the stereotypical picture of an unqualified buyer who might be subject to a “predatory” loan during a credit boom. Indeed, the aggregate number of homebuyers from the observed cities in the 2nd and 3rd income quintiles actually decreased from 2003 to 2005 (the boom peak), while high-income homebuyers increased.

This is the story of a broken housing supply reshuffling the deck, not of irresponsible lending to marginal households.

#### IV.

Even if most people could be convinced that there wasn't a classic bubble in housing, it's an even taller order to convince them that the larger financial and economic crisis wasn't a direct effect of the housing boom-and-bust.

Erdmann devotes much less time to the connection between the crises, but I think it would have been better to dive in those deeper waters (though, in fairness, in doing so, he would have been leaving behind his original contribution and getting into topics that are contentious, and probably irresolvable, among professional economists). He later did so [in a paper with Scott Sumner](#) laying out the thesis, and I find it just as convincing.

But, even in *Shut Out*, he provides enough evidence to be convincing in its own right. It shouldn't take much, because if there was no bubble, there didn't have to be a crash; and it's entirely plausible that whatever caused the housing crash could spread to the larger economy. Time and time again, the federal reserve, politicians, academic economists, and the financial media hammered home the message that there was an irrationally overheated housing market, the correction was coming, and it was a matter of just when and how bad.

Markets are forward-looking and all relevant data matter, and the housing market is no different. From 2004 until 2006, two things happened - the federal reserve had begun raising rates and Bernanke et al. had begun cheering on declining home prices and construction. It's no surprise, then, that is exactly what happened. The market could see the writing on the wall. There was an ensuing increase in mortgage defaults due to rising prices and higher interest rate resets.

But this was likely manageable and would not have been a big deal by itself for mortgage backed securities and CDOs; the problem that led to a panic is that *policy makers showed no signs that this was going to stop because they had assumed it was all necessary*. Once that panic set in, and because the Fed was so laser-focused on housing, they missed many economic signals of a pending recession (most notably, collapsing projected nominal GDP), with a disastrous series of decisions to follow.

So, no, the housing bust did not cause a financial crisis that led to a recession. It is more accurate to say that the same forces - bad monetary policies - caused all three. And then for the cherry on top: post-recession legislation severely tightened credit markets which persist to this day. This turned a deep recession into a deep recession plus a long, slow recovery, as the housing and credit markets in most of the US still haven't returned to 2000-levels. Homeownership rates for Generation X and younger in 2017 were the lowest they've been since before 1982. Housing prices in low-income areas outside of major cities, which did not participate in the price boom, nonetheless saw the largest decreases in the decade to follow.

Erdmann: "There is an awkwardness...in the credit bubble theory...which was solved by invoking behavioral finance. Buyers and lenders were irrational, and a collapse was unavoidable... A model whose very specifications are that buyers and sellers are unmoored from objective reality is difficult to falsify. So, as home prices rose higher and higher in 2004 and 2005, even as the Fed raised...rates, the model was adjusted to account for buyers and leaders who must have been even more irrational than previously thought...The problem is that, if a behavioral model influences policy, then collapse is imminent *whether the model is right or wrong* - if not because of agent irrationality, then because of model error. Those who 'called' the bubble have regarded the collapse as confirmation that there was a string or irrational players... **What if model error in public policy was the primary cause of the collapse?**" Indeed. (Emphasis mine.)

(None of this is to defend the risk models of investment banks. They clearly were off in a specific way. The underlying premise appeared to be that a nationwide collapse in housing prices was very unlikely, given the local nature of real estate prices, and the securities were geographically diversified. I actually agree with this premise; but one has to account for low probability correlated risk, even if that risk is that the Fed will screw up. Still, it would be worth it to come to a broad realization that the more fundamental problem was policy error.)

V.

Most of the latter half of the book discusses the specific broader economic effects of our implicit policy to wall off centers of growth and opportunity from a majority of the population, *and to also chase away the people that are already there to begin with*. This part is not particularly original, and is likely well-trodden ground for readers of this blog. Nonetheless, it's important to shout from the rooftops - rinse, repeat - until the truth becomes unavoidable. If we are going to develop a collective narrative about housing, it might as well have the virtue of being correct.

“Closed access housing has become such a high barrier to the free movement of capital to its most valued use...that the US population appears to be fleeing prosperity. The effect this is having on our national standard of living - especially for households with lower incomes - must be significant... [Data] shows that rising income is associated with migration away from a city. We should surely expect migration patterns in a liberal society to operate in the opposite direction... But through local housing constraints, we have created an economy that is shaped by exclusion... income is determined by access, and households are pushed away from affluence. Exclusionary housing policies are the root cause of wage stagnation, asset inflation, income inequality, and the “bubble” economy.”

Erdmann presents graph after graph that confirms the story of low-income residents of high income cities fleeing to lower-income areas, leaving them with more to consume net of housing costs. (More to consume, that is, of material goods; one can only assume the consumption of the non-material good of living in an amenity-rich dynamic city is worth something to these people, or they would have left before being squeezed out.) And as the proportion of the population in high-income cities has shrunk relative to those living in lower-income cities, measured inequality has grown.

The upshot is that economic innovation has slowed down as the barriers to highly innovating locales have been maintained; consumer prices are higher than they would otherwise because of what is necessary to compensate landowners in closed access cities; and economic opportunities are unavailable to the poorer among us. All because of short-sighted local housing policy in a handful of places.

(This was written in 2018; it remains to be seen how the post-COVID economy affects these dynamics. It is plausible that these could rapidly reverse independent of change in policy in these areas, as work-from-home and an extension of the housing-related exodus into higher income brackets could shock the supposed network effects of highly innovative cities, causing a cascading reversal of this trend).

He even shows the effect this had had on international trade. Erdmann (smartly) steers clear of discussion of political party dynamics, but it cannot be a coincidence that the era of the housing boom-and-bust, with different dynamics playing out in different geographies, just so happened to occur along with the acceleration of a great reorganization of political alliances. While problems in education and health care are not due to real estate costs (though, do you know how much your doctor is paying in rent?!), similar dynamics are at play where, in the words of economist Arnold Kling, we effectively “subsidize demand and restrict supply”, and then are surprised by the outcome. The lessons of housing do generalize.

I’m not saying everything is explained by housing costs; I’m just saying whatever issue you’re taking up, see how much, if any, is explained by this dynamic and then see what you have left.

## VI.

The only real negative of *Shut Out* is the necessary slog through a data-rich jungle of information; many of the data figures are great in navigating the terrain, but they can be difficult to read, and some colorized charts would have been very helpful. I admire the author for forging ahead anyway; part of the problem that made the actual dynamics at play so difficult to recognize in real time was the reliance on aggregated national-level data that told the opposite (and false) story from the one elucidated by the disaggregated local-level data in *Shut Out*.

The reward for doing so is in the epilogue, where Erdmann sings a beautiful hymn to an open and dynamic economy. He talks about the open access orders that have evolved in the developed world (especially in the Anglosphere), in contrast to the closed access orders that persist in the developing world, where opportunity is limited in economies where access is a political favor to be granted.

The housing markets in the richest cities in the US are becoming more like developing world economies in the way that access is granted by either paying unnaturally large sums (for new renters and owners) or by navigating a thicket of economic rents in the form of taxes, fees, and political favors (for developers). Even worse, with a better developed formal legal and political system, enforcement of this closed order is more effective than what you see in the poorer countries. Navigating the closed access order in Brazil or India may be as simple as paying a petty bribe; in S.F., to build new housing you have to pay large fees and convince power brokers at public meetings, and you might still get shut out of the process (but no refunds!).

In the end, the very future of our economic system may hang in the balance; because of the effects of housing policy, and our inability to properly diagnose the downstream problems, “there is a growing sense that spoils of growth are only shared by a select few. The politically imposed limits on how many people can share in prosperity undermines the communal support for growth.” Erdmann seems pessimistic that this can be solved in these markets, given the entrenched interests, and offers the history of decline in Detroit (for reasons other than housing) as a warning. But the rest of America and the cities that might challenge for economic supremacy may be able to learn from the mistakes of the closed access cities, and elect to build liberally. His advice: “let it rip.”

## VII.

In college basketball, former UNLV coach Jerry Tarkanian, an outlaw who skirted the rules of amateurism and who was a constant critic of the institutional framework of his sport, once famously quipped: “The NCAA is so mad at Kentucky it will probably slap another two years probation on Cleveland State.” His point was that the NCAA protected larger cash-cow programs and saved their enforcement firepower for the also-rans.

It was probably never quite true of the NCAA, but if the ghost of Tarkanian came back to say “the federal government got so mad at San Francisco for not building enough housing that they engineered a recession and forced Ohio homeowners to lose their homes”, I’d say it was the most accurate (and fair) description of the late 2000’s that had been uttered.

Indeed, if the engineered housing bust and recession had somehow solved the housing supply crisis in superstar cities, I suppose it would be worthwhile to debate whether it was worth it, even if there were other better ways to solve the problem. But it solved nothing. More than 10 years on, the cost of housing and rents in these locales is as high as they were during the boom years (temporarily falling rents due to COVID notwithstanding), while home prices in the rest of America still haven’t recovered and potentially willing and able homeowners still can’t qualify for a mortgage.

But I doubt any of these problems can be solved without a reckoning of how most everyone got the narrative wrong, and how the imposed solutions likely made things worse in the long run. Hopefully, *Shut Out* is the beginning of that reckoning.



# SPQR by Mary Beard

SPQR is the captivating account of how a nameless Italian village became one of humanity's most influential empires, the impact of which dramatically shapes our modern lives without most of us even realizing it.

The Roman Senate and People – the “Senatus Populusque Romanus” for which the book is named – rose up around 750 BC from their mythical twin founders, Romulus and Remus, and grew over a period of centuries into one of the grandest republics of the ancient world.

You might think 608 hardcover pages of ancient history would be a formula for curing insomnia... but you'd be wrong. Mary Beard mesmerizes with her well-paced retelling of ancient Roman history, deftly weaving first-hand accounts, recent archaeological discoveries, and multiple historical sources into a tale that leaves the reader with a profound appreciation for the many things wrong with our popular understanding of Roman history as well as the pervasive influence that ancient Rome continues to exert on our modern world.

Why do our days start at midnight? (The Romans started this.) Why are the liberal arts called “liberal?” (They were studies meant for the “free” people of Rome). Where do “July” and “August” come from? (Julius Caesar and Augustus). The book is filled with context-setting examples like these.

Long before it was the familiar empire of Shakespeare and numerous other popular dramatizations, Rome was a republic ruled by pairs of Consuls – the most famous of whom was the eloquent Cicero – and a Senate that voted on the most important decisions for its populace. Ancient Rome was in ways more progressive than is commonly understood. For instance, its Senators came from throughout the republic, representing the diversity and origins of even conquered peoples. Slaves could (and did) regularly earn their freedom, becoming citizens indistinguishable from any other free Roman. People often traveled vast distances in the empire, intermingling and settling in new locations. For instance, Quintus Lollius Urbicus, ruler of London in the years 139-142, was Algerian. The most eloquent words spoken against Roman rule were often ascribed to the mouths of conquered peoples by sympathetic Roman writers. However, ancient Rome ran in other ways opposite of modern times. The poorest lived on the topmost floors of buildings while the rich lived on the bottom floors, preferring stair-free access to the street. The rich were regularly served food at home while the poor ate out because they couldn't afford the space and equipment.

The book also contains amusing facts. For instance, Augustus made himself the most common surviving image of any person from the ancient world by putting himself on coins throughout the empire. And interestingly, there's an inexplicable change from



universally beardless visages to beards after Trajan (117 AD), which remains a sure-fire way to identify the relative age of coins you find from ancient Rome.

SPQR starts from the foundation of Rome around 750 BC and ends at 212 AD, when the Antonine Constitution was issued by Emperor Caracalla, granting universal citizenship to all free men in the Roman Empire.

Much has been said and written about the fall of the Rome, but the story of its rise is arguably even more important, especially when it comes to its relevance to the modern world. SPQR belongs on a very short list of must-read books when it comes to the rise of empires. I highly recommend it if you're at all interested in history.

(I listened to this over a handful of hours using the Audible version, read expertly by Phyllida Nash. Also recommended.)

Veni, vidi, vici (this book). And you can too.

# Steve Jobs by Walter Isaacson

I.

If you ask most people what makes them a good person, generally they will give you some estimation of how nice they are to others. They will talk about how they recycle, or how they held the door for an old woman, or how they engage in similar acts of kindness throughout the day. For most people, being the best they can be and being nice are the same thing.

Enter Steve Jobs.

Most people think of Steve Jobs as either a petulant man-child who demanded his own way in even the most pointless circumstances, or as a visionary whose ability to understand how regular people use technology in their daily lives catapulted the entire world into the Information Age. Walter Isaacson's definitive biography of his life, *Steve Jobs*, portrays a man who was simultaneously both. Despite being released very soon after Jobs's death, the book is nonetheless an extremely relevant read today, almost a decade later.

I think the reason for this is because the book is a high-fidelity representation of a man whose inability to even basically progress on some social levels was not a serious obstacle for him to make a positive impact on the world. Despite many stories in the book about Jobs's frequent tantrums ("‘Let's stop this bullshit!’ he kept shouting"), I nonetheless walked away thinking that on the whole, the world is a better place because of Jobs's existence. That's in part due to Isaacson's extremely measured style of writing. He alternates between anecdotes about Jobs's immaturity and vision with relative ease. There are few points in the book where the reader cannot both empathize with Jobs and yet feel kind of vaguely put off by him. In this way, Isaacson puts the reader in the place of those closest to Jobs: kind of like a tour of his famous reality distortion field. It's a little jarring at first, but that's hardly Isaacson's fault. Job's life was riddled with contradictions.

One might ask how much of the book is, in fact, a clear representation of Jobs's life, due to the aforementioned distortion field. Reassuringly, Isaacson presents a list of his efforts to be objective in the book's opening, including interviews with figures who had much reason to love or hate Jobs.

To check and flesh out his story, I interviewed more than a hundred friends, relatives, competitors, adversaries, and colleagues.

These people include Bill Gates, Laurene Powell, early Apple engineers, Jony Ive, and more. Isaacson also indicates the possibility of the [Rashomon Effect](#) taking place in these interviews. All of this contributes to the air of mystery and awe around Jobs- just who was this man who inspired such love, hate, and adulation?

## II.

The book is impressive because it somehow answers this question. Despite being fairly short at less than 700 pages, it nonetheless somehow gives a nuanced portrayal of Jobs's life and steers away from leaning too heavily to one side in the "visionary vs man-child" debate. At the end, Isaacson briefly gives his own opinion of Jobs's life. This feels like catharsis because, during the few times when Isaacson editorializes before this, it's almost always to disprove some "fact" that Jobs is claiming or provide context to something that Jobs was avoiding. So when Isaacson does give his opinion, it feels very earned. Two quotes from this section stand out to me in particular:

Was he smart? No, not exceptionally. Instead, he was a genius.

This may feel like puffing up of Jobs. In the context of the book that precedes this quote, nothing could be more wrong. A genius, as portrayed in this book, is someone who is truly world-class at something, but fundamentally flawed. Jobs could not have been who he was if he had fixed his underlying problems. Which leads me to the next quote:

Most people have a regulator between their mind and their mouth that modulates their brutish sentiments and spikiest impulses. Not Jobs. He made a point of being brutally honest. "My job is to say when something sucks rather than sugarcoat it," he said. This made his charismatic and inspiring, yet also, to use the technical term, an asshole at times.

And he truly was. At times, Jobs comes across as shockingly cruel. There are anecdotes about him refusing to grant founder stock to a co-founder for no apparent reason. His regular habit of parking in handicapped spots both amused and annoyed his co-workers. And, perhaps most egregiously, he refused to acknowledge his own daughter's existence for several years.

And yet, by the end of the book I'm wishing that Jobs was still around, reinventing one more industry or even just making one more device. How can these facts be reconciled?

## III.

I think the answer lies in the fact that Jobs was kind of a child in several ways. For some reason, he never lost the wonder and curiosity that is almost ubiquitous to small children. And he also never lost the cruelty that children often display. To this day some of the most abjectly mean things I've ever heard have come from children. They simply have no filter. Neither did Jobs. This would also explain his frequent crying fits when things did not go his way. He was simply reacting emotionally, as all children do.

This is an interesting idea to me, and the book doesn't really dive into it the way I might expect it to. Again, the book is short, and again, Isaacson isn't much for editorializing, so fair enough, I guess. But there are so many stories of people overcoming struggles by growing up, by learning to follow

rules, by becoming socialized. In fiction, this usually takes the form of the Hero's Journey. However, Jobs does none of these things and still succeeds. In a huge way. That's not to say he doesn't grow at all throughout the book. Jobs's initial run at Apple can be contrasted with his later return. But while he did improve, he never really got to the level of a normal person. I don't think it's a stretch to say that, were he still alive today, he still wouldn't be at that level. But Isaacson makes me ask if I would really want him to.

Does this mean that the Hero's Journey is baseless and real life is a lot less optimistic? I don't think so. I still think that the average person should probably try to pick up some social skills and maybe work on their LinkedIn or something. It's just interesting to read about someone who doesn't do that kind of thing, and moreover, someone for whom that kind of thing would not have worked. I think the Hero's Journey has such broad appeal because many find it's a great template towards self-improvement. This leads us to forget that there are some people for whom that might not be the best path. I won't be parking in any handicapped spots anytime soon, but I do think it's valuable that the story of someone like that exists. It's a testament to Isaacson's work that I hope the next Steve Jobs reads this book.

This opens the door to a concerning question- which rules should really apply to the truly exceptional in our society? Obviously there are some rules which everyone should follow- murder is probably wrong regardless of if you're about to invent the smartphone. But if we had to give up handicapped parking spaces to get a few more Steve Jobses in the world, would that be the worst trade?

This postulation fails to take one crucial thing into account, though. If it wasn't against the rules to park there, Jobs probably wouldn't have done it. He absolutely enjoyed being a contrarian and rebel. Almost all of his work incorporates punk and hacker ideology, at least on the surface level. So even if a mystical and perfect cost/benefit analysis showed that a net positive happens when the exceptional can park in handicapped spaces, they'd probably just start reckless driving instead.

I think this discussion is much broader than this book review or even this book. But *Steve Jobs* did open my mind to that possibility. I guess even after Jobs was gone, he was still making people [think different](#).

# Strangers Drowning: Impossible Idealism, Drastic Choices, and the Urge to Help by Larissa MacFarquhar

## I.

Peter Singer famously argued in *Famine, Affluence and Morality* that “if it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it.” He illustrates this idea with the drowning child thought experiment: if we see a child drowning, and it is in our power to save them at an insignificant cost to us, then it is our moral duty to do so. It seems most people agree in the case of the drowning child, even when the monetary cost of the damage is explicitly mentioned. When it is pointed out that the situation of the drowning child is similar to that of children at risk of dying from malaria however, people recoil, although they don’t necessarily claim to be able to express why those situations would be different.

It seems clear that this difference comes from emotional biases: our moral intuitions are difficult to scale up, we care more about an identifiable victim than about a large group of anonymous people, the bystander effect lets us think that somebody is going to intervene (although the fact the people are still dying from malaria makes it clear that our help would be far from useless), thus [we care more about people physically close to us than people far away](#). We are also programmed to care more about people emotionally close to us than strangers; we could discuss [whether this is a good thing or not](#), but until we have the means to change our biology, this is a fact of life, and we have to take it into account. But the fact that we care more about friends than strangers doesn’t tell us where we should draw the line between friends or strangers, or how much we should help people depending on their emotional distance to us. Some people are able to keep in mind the death and suffering that their lack of action would represent, and keep acting accordingly to help those in need. In *Strangers Drowning: Impossible Idealism, Drastic Choices, and the Urge to Help*, Larissa MacFarquhar writes about such people.

The first chapter explains what a do-gooder is. Simply put, they’re “people who try, more than anything else, to do good.” This implies that they organise their lives around moral principles, to live as ethical a life as possible: a do-gooder is “someone who pushes himself to moral extremity, who commits himself wholly, beyond what seems reasonable.” This commitment comes from a strong sense of duty: several of the do-gooders in that book say they have no choice, and see their actions as something they are duty-bound to do. They see everything wrong in the world as their problem: knowing is enough to make action a moral necessity. They may have personal dreams and desires, but those come second to fulfilling their duty, or at least coming close enough to having done so. But for many of them, helping is already an end in itself, and they

argue that we should be [Giving Gladly](#), as one of the characters named her blog. Some of the do-gooders find a real joy in having a useful purpose. But conversely, being happy makes you do better work. The do-gooders certainly are on a continuum with the average person. MacFarquhar gives the example of war:

“In wartime—or in a crisis so devastating that it resembles war, such as an earthquake or a hurricane—duty expands far beyond its peacetime boundaries. In wartime, it’s thought dutiful rather than unnatural to leave your family for the sake of a cause. In wartime, the line between family and strangers grows faint, as the duty to one’s own enlarges to encompass all the people who are on the same side [...] This is the difference between do-gooders and ordinary people: for do-gooders, it is always wartime. They always feel themselves responsible for strangers—they always feel that strangers, like compatriots in war, are their own people. They know that there are always those as urgently in need as the victims of battle, and they consider themselves conscripted by duty.”

Most of the chapters of the book discuss the life of one of those “do-gooders” (although they often attract and follow other people with similar commitments). Their similarities are as interesting as their differences (although we’re relying on MacFarquhar’s narrative here; it’s possible that she decided to insist upon certain aspects to give them more importance than the characters would grant themselves). The most obvious difference is the diversity of causes: humanitarian aid for Dorothy Granada, Kimberly Brown-Whales and Baba Amte and his family; Effective Altruism for Julia Wise, Jeff Kauffman, Stephanie Wykstra and Geoff Anders; animal welfare for Aaron Pitkin (not his real name); suicide prevention for Nemoto Ittetsu; international adoption for the Badeau family. The choice of a cause is an important decision, as it is not possible for one person to right every wrong in the world (except for Geoff Anders, who has decided to do precisely that. Good luck to him!). MacFarquhar quotes E. M. Forster here:

“And why should the divine hospitality cease here? Consider, with all reverence, the monkeys. May there not be a mansion for the monkeys also? Old Mr. Graysford said No, but young Mr. Sorley, who was advanced, said Yes; he saw no reason why monkeys should not have their collateral share of bliss, and he had sympathetic discussions about them with his Hindu friends. And the jackals? Jackals were indeed less to Mr. Sorley’s mind, but he admitted that the mercy of God, being infinite, may well embrace all mammals. And the wasps? He became uneasy during the descent to wasps, and was apt to change the conversation. And oranges, cactuses, crystals and mud? and the bacteria inside Mr. Sorley? No, no, this is going too far. We must exclude someone from our gathering, or we shall be left with nothing.”

The sheer size of the mission do-gooders take up would be a recipe for immediate burnout if they were to throw themselves at it impulsively. Many of the do-gooders presented in the book start somewhat haphazardly following their impulse to do good, then come to realise there’s more to do and they need to find better methods, so they meet other activists and read about the topic (Aaron’s approach is particularly interesting: early in his activism, he is regularly reading anti-animal rights literature to find any compelling argument that could show his quest to be misguided; after a few years of not finding anything, he decides he’s clearly right in wanting to improve chicken welfare). Many of them also read moral philosophy: Peter Singer is mentioned as a

direct inspiration in several cases for instance. But even for utilitarians, rules remain an important part of their framework, even to such extremes that they stop producing results. One example is how Julia sees any purchase as a waste of money that could have bought anti-malaria bed nets instead; but this reasoning fails to take into account how the stress this places on her can decrease her effectiveness in her activism. The solution also comes from rules:

“Jeff decided he would give away 50 percent of his salary and keep the rest for spending and saving; Julia would give away 100 percent of hers. Out of the remainder of Jeff’s salary, he allotted an allowance to each of them of thirty-eight dollars a week, which they would use to pay for everything other than rent and food—things such as clothes, shoes, transportation, and treats like candy apples. Jeff decreed that this allowance had to be spent on these things: it could not be given away, and it could not be saved, or he would donate a matching amount to the Republican Party. That way, if Julia wanted to spend money on something, she would not be taking that money away from someone who was dying. (Julia realized, of course, that this wall they had set up between his money and hers existed only in their heads, but since its only function was to preserve her sanity, that didn’t matter.)”

Aaron’s commitment to his principles can be quite myopic: he claims for instance that any time spent on doing the dishes would be time spent not working on saving animals; considering his devotion to his work, this may be true, but is he taking into account how the stress the mess places on his girlfriend impacts the productivity of them both? He refuses to lend his girlfriend the money to cancel her credit card debt, because most of the money he has comes from his grandfather making fur coats, and thus should only be spent on the worst sort of suffering; this ignores the fact that any money paid as interest payments to credit card companies is not sent to charities. He also rejects marriage because it was still illegal for same sex couples in Massachusetts: I fail to see the impact on his individual decision on this legal question. Later one of his girlfriends, who is similarly extreme, breaks down in tears when he offers her a second-hand DVD player, because it might have been made by exploited children; but since it’s secondhand, I really don’t see what possible impact their purchase could have. Even someone who sees their action as completely rational can have reactions that only make sense from an emotional perspective; this emotionality becomes a hindrance when it comes to effectiveness, and do-gooders have to learn to become more detached to fulfill their mission. Aaron finally learned to put caveats in his principles when it really matters, such as eating meat when meeting with cattlemen, as being approachable can do more for his cause than refusing that one steak.

The clearest example of a do-gooder failing to see the consequences of his actions is Charles Gray. We’re introduced to Charles in the chapter on Dorothy (he was her husband). Charles’ big idea is the “World Equity Budget”: wealth inequality is a big problem, and everybody should live on their fair share of the world’s wealth (or even less than that, to take future generations into account). Although he keeps reevaluating the amount, the numbers we’re given are \$62 a month at one point, and \$1,200 a year at another, to live in the USA. Charles is a bit of a running joke in the book, as each of the do-gooders who has heard of him seems to think he’s crazy. Even Aaron finds the WEB a bit extreme, and Julia is saddened that such a strong moral impulse gets wasted

on something so pointless: why would it be more important that he be poor than that other people be less so? Even Dorothy, who joined him in all of his insane projects:

“I told Charles Gray, ‘Okay, I’m in love with you. I love the project. I’ll join the World Equity Budget. But I must have my glass of wine before dinner, and my coffee. I will not do without that.’ We used to fight—he was always adding up every penny, he kept a notebook. Once, he told me that I was thirty-eight cents over budget. I said, ‘Would you repeat what you just said?’ And then I told him what he could do with his World Equity Budget.”

Apart from Charles Gray, each of the do-gooders seems to agree that rules are useful tools, but are less important than the goal they serve.

Another set of rules is heavily discussed in the book: religion. In some cases, the reason is obvious: Baba, Kimberly and Ittetsu have all been clergypeople of their respective religions. The Badeaus prayed on it every time they had to decide whether to adopt a given child (or siblings). Stephanie’s story is mostly focused on her crises of faith. But in the case of Julia and Dorothy, their progressive loss of faith does not seem as central to the story. In the case of Aaron, a nonpracticing Jew, MacFarquhar judges it useful to point out that he disliked religion. One possible reason for this insistence on religion could be the historical ties in many cultures between religion and charitable giving. Typically, it used to be considered normal to give 10% of one’s income to the church; that’s one of the reasons Toby Ord chose to recommend this threshold for [Giving What We Can](#). Another reason of MacFarquhar’s focus on religion could be to highlight the religiosity of doing good: throughout the book, the do-gooders are compared to saints (and in the case of Baba, MacFarquhar claims he “read the life of Christ not as an account of perfection but as a practical template”), and they themselves feel a connection to the divine through their charity: Dorothy’s fast against nuclear weapons is a political act, but feels like a kind of prayer and a communion with God; the act of giving a kidney to a stranger is described as “a transcendent experience, akin to a religious one;” Ittetsu’s counselling of suicidal people is informed by the training in Zen listening he received as a monk; Stephanie’s faith in utilitarianism is a replacement for a broken faith in Christianity. MacFarquhar isn’t denying the obstacles that religion can pose to doing good though: one of the rare mentions of religion in Aaron’s chapter is the description of kosher beef-slaughtering, which is presented as one of the worst abuses that we inflict on animals; both Dorothy and Kimberly are opposed in their attempts to help women on the basis of Christian morality; Baba and his family become outcasts for daring to help the lower class “Dalits” (also known as untouchables).

Other similarities frequently highlighted in the backgrounds of the do-gooders are, on the one hand their caring and generosity, and on the other hand, the hardship and abuse they had to go through. The first part is made up of cute stories about do-gooders bringing home stray animals and giving away their allowances, as well as heroic ones about rushing toward disaster areas or fighting soldiers harassing a woman (I need to find a Bollywood flick about the life of Baba Amte, there’s incredible potential in there!). Including these stories makes sense when we’re trying to understand an extreme impulse to do good. The second part includes abuse, depression, growing up in an



unsafe and unstable environment, alcoholic parents, sick relatives, suicide... Here, I'm not sure whether these elements are mentioned because of how they would impact someone, or if MacFarquhar is making a statement about how suffering could be a key to understanding do-gooders, either because enduring this suffering prepares them for the horrors to come, or because suffering can be the mirror through which they learn to see the cracks in the world and the reason they decide to make it their life's work to close some of those cracks.

Indeed, the best way to do good is not necessarily in the field, and a strong sense of empathy is essential to maintain one's commitment when the people (or chickens) being helped are not even stories but numbers. Some people are better than others at seeing the forest for the trees, and when Aaron needs to feel that "warm glow" of directly helping others, it's enough for him to be carrying boxes and cooking the meals and let other volunteers serve them and get the direct thanks. It's for a similar reason that he chose to focus on chickens rather than cats and dogs, who already had people caring enough to protect them: the more neglected the cause, the better the return on investment. This ethos is at the heart of the effective altruism movement. Julia and Jeff share this utilitarian perspective, and once their rules for giving were set up, they could simply enjoy knowing what their money was doing, without having to see it directly, except through GiveWell's evaluations. Still, that joy can never become the main purpose of giving, and it's more important to think about what people need, rather than what it feels good to give: "What was important was that people were helped, not that it was she who was doing the helping." Julia comes to the realisation that she should be earning to give: the goal being to maximise one's impact rather than to give as much as possible of her money away, it would be more efficient to earn a higher income. She doesn't have to go for the absolute maximum income accessible if that would mean being so miserable it would end up with a breakdown, but surely there must be a better position? She considered psychiatry for a moment, and the combined effect of the high income and her blog encouraging people to give more would certainly have resulted in a high impact (although... a psychiatrist with a blog? That would never work). In the end, she decided to remain a social worker.

After all, there are other ways to maximise one's impact. One is to encourage others to give. But there's a fine line between not doing enough and being so preachy you put people off. Kimberly mentions having to restrain herself when encouraging her parishioners to give, for fear of receiving nothing. For an example of this done well, we have the example of Toby Ord, who came to the conclusion that starting Giving What We Can had been the most effective thing he could ever have done, far beyond the impact of his personal donations. MacFarquhar also mentions two students of Peter Singer whom his teaching convinced to give in their first year out of college more money than Singer probably ever made from writing or teaching. Knowing how to maximise one's impact is a hard question: donating one's money to a [GiveWell](#) recommended charity certainly is a good thing, but that's the safe option: when thinking about her options for a career, Stephanie considers life extension and AI alignment ([MIRI](#) is directly mentioned). And although Julia relates an EA dinner where people with more traditional approaches to charity were dismissed, people are still needed on the ground, and this is not an easy job.

Being a do-gooder means being constantly aware of the suffering everywhere, and often witnessing it yourself. You will even have to share some of it. Helping people who live in unsanitary conditions means having to see a lot of death, always wondering what more you could have done (plus all the diseases that you will catch). Working in war zones means you could be kidnapped or raped. Building a village in the wilderness means that until it's built, you'll have to survive everything that's trying to kill you. Helping others means being made painfully aware of [how bad things can be](#).

Toward the end of the book, MacFarquhar insists that the only reason she's writing about these people is that they were lucky enough not to die:

"Dorothy Granada did not fast to the death and was not shot by Contra soldiers. Aaron Pitkin did not remain homeless or alone. Julia Wise did not stay childless. Baba Amte did not contract leprosy, and his children were not eaten by panthers. Prakash Amte was not forced to cut his baby to pieces to save his wife, and was not killed by one of his animals. Kimberly Brown-Whale's son did not die in Mozambique of his heart condition, and her daughter was not kidnapped. Ittetsu Nemoto did not die from overwork. Sue and Hector Badeau held their family together. But any of these stories might have turned out differently. One step over the brink and their commitment might have looked like craziness, or cruelty."

And of course, helping means accepting all those facts and going on, despite the fear, when the do-gooder could simply walk away. This sense of injustice is very present when women in Baba's family have to choose between giving birth in unsanitary conditions, or leaving the people who need their help because they don't have that same luxury of leaving. And then later, is it fair to her child to come back with them, when they don't have to be there and could enjoy a safe and comfortable life? But why would she enjoy this privilege when all the women around her are likely to lose some of their children? (after failing to deliver a baby, Prakash Amte is told that "many babies died in their first year of life from malnutrition or illness or snake bites; if a mother died in childbirth, the father usually didn't take care of the infant—it was left to die. This death was just a little earlier, that was all.")

Whether doing good through donations or in the field, the decision of organising one's whole life around one's moral principles is going to have an impact on those who share that life. There is self-selection at work however: since do-gooders' social circles tend to be mostly charity workers, activists and generous people, the partner of a do-gooder is more likely to share their values to some extent. It's hard to guess how much beforehand though: Aaron's girlfriend Jen saw herself as committed to their shared values, but ended up leaving him, being dissatisfied with the relationship. What hurt her was that she had to admit that she was leaving him because she felt she was not as ethical a person as he was: "she was not just leaving Aaron; she was choosing selfishness. She was choosing her own happiness over the survival of other creatures. She could not justify it, she thought it was the wrong thing to do, but she couldn't help herself; she wasn't him." The same thing happened with Charles' first wife, who despite giving half of her money already, refused to follow him with his WEB project. Dorothy accepted the WEB, but even she had a similar problem: "We're so serious, all us peace-and-justice people. These people, they can't have fun, they have to be out on the firing

line all the time.” “Those people did wonderful work, but they were really not nice people,” “They were people you did not want to be around. They were so sharp. Everything was a matter of life and death: we’ve got to do this action because the world depends on it.”

In other cases, the partner fully buys into the project, like Jeff with EA, or Hector with his wife's project of adopting children (after the fourth child, it was Hector who was pushing Sue to adopt more, up until their 22nd child). An interesting case is Baba's wife Indu, who married him when he was a rich Brahmin. Because she was taught to obey her husband in all matters, she follows him when he decides to work with Dalits, accepting her own reputation to be tainted too. She remains by his side when both of their families reject them, and when he renounces his wealth, and even when he starts working at a leprosy clinic, and again when he starts his own ashram of Anandwan (a shelter for lepers made of wooden sticks and grass, to live among rats, tigers, leopards, under 115 degrees with no water. The name means “Forest of Joy”). Indu stays and takes up a large amount of the workload. She takes the name Tai, sister to the lepers. She takes care of everyone there, making little distinction between the patients and her family. She only seriously complains to Baba after he's left Anandwan to join the anti-dam movement on the Narmada river. Tai feels useless there, and is miserable; after ten years, Baba accepts to return with her to Anandwan. Similarly, their children and grandchildren join the family mission, with one son improving the ashram, and the other creating another clinic nearby to take care of the native Madia Gonds, as well as an orphanage for injured wild animals. By contrast, the Badeaus children didn't always understand why their parents were adopting so many children (after a bike accident, one of their daughters complained: “If you didn't adopt so many children, I could have had a new bike, and this wouldn't have happened to me!”). In the end, none of the 22 children decided to adopt themselves.

Ittetsu is one rare case for whom his partner does not seem to have been involved in his charity in any way (or at least that's the way she appears in MacFarquhar's book; the [documentary on his life](#) seems to show her as very supportive of her husband's work). He's also the one who suffered the most severe burnout. There are several reasons for that. The first one is the extreme empathy he shows to suicidal people, which takes a toll: while he was maintaining his blog (“For Those Who Do Not Want to Die”), he was making the effort of replying immediately to everyone in need of counselling, day or night (the messages mostly arrived at night), just a bottomless chasm of anxiety threatening to engulf him. This is a common problem for do-gooders: “any do-gooder who is not dead or irredeemably jaundiced by the age of thirty has learned to acquire a degree of blindness in order to get by.”

The second reason is the apparent lack of progress, sometimes for years: there simply wasn't much he could do beside listening. He develops coping strategies, but they aren't enough. Five years within his practice, he develops an unstable angina and had heart surgery four times over the next two years. He has to stop counselling for a few weeks, and explains his situation on his blog while in the hospital. Of course he receives messages of support... and other kinds too: “They didn't care that he was sick: they were sick, too, they said; they were in pain, and he had to take care of them.”

Other do-gooders reported similar reactions: Dorothy and Charles made other activists angry when they gave up on their fast against nuclear weapons, saying that you can't abandon a fast to the death. When they leave the movement, they are criticised again for abandoning the higher calling for peace to selfishly treat injured people in a war zone (in fact three weeks later, [the world came close to its end](#). Coincidence? yes). On the other hand, protesting nuclear weapons got Dorothy this reaction from her black friends: "What are you doing with that white movement? Let them blow themselves up! To hell with it! Who cares?". Because nuclear winter can only affect white neighbourhoods of course. Basically, if you decide to start working for a cause and switch to a different one, you're a traitor, as Baba can attest following his decision to leave Anandwan to prevent the building of the Narmada dam. But if you keep fighting for that cause, you need to be able to pass whatever purity test the other activists have decided for you. One of the reasons "Aaron" uses a pseudonym in his chapter is that he's afraid of what could happen to him if other activists were to learn he sometimes eat meat to achieve better results:

"He thinks about the example of a famous animal-rights activist who he'd heard was an awful person with a coke habit, given to sexual harassment and waving guns about, but the thing that really scandalized his followers was the rumor that he'd been spotted eating M&M's. True, the people who raised hell about the M&M's were a bit crazy, but they were Aaron's base; he needed them to do his work."

If you're working intensely to achieve next to zero result, and all you get as a reward is a deteriorating health and abuse, is this really worth it?

## II.

*Strangers Drowning* aims to answer two questions: the first, why people go through such lengths to help strangers is answered through the biographical chapters. The second is why the rest of the world doesn't help more. The non-biographical chapters deal with various topics, but the general theme is the world's reaction to do-gooders.

So why don't people do more to help strangers? Regarding giving, loss aversion is one explanation: who knows what crisis looms in the future, that might require just the amount of money we've saved so far? What would happen then if we have given away a large part of our finances? But we're not just saving the money, we're also spending it on things that are objectively less important than the lives we could save instead. And here, it all comes down to our lives, and those of our relatives, being more important to us. Another important factor is the identifiable victim effect: writing checks is "a dull way of giving", without glamour; it would be more motivating to see direct results of our charity, and contributing mere drops to extinguishing the inferno of misery is too daunting a task (and our intuitions are bad at scaling far above human scales). As a result, it's easier to simply avoid the cognitive dissonance by not giving at all (or just enough to avoid the opprobrium of not giving). Even when people give, Toby Ord laments that it's hard to get them to give more effectively, because of change aversion.

Still, McFarquhar admits that “Should most of us do more for others than we do now?” is not a hard question. But for people who are giving a decent enough portion of their wealth and/or time to see themselves as good people, and are aware of the warm glow we feel when helping others, why can’t we be do-gooders too?

Here, it comes down to the sacrifices that come from being a do-gooder. Some of the sacrifices are small, so small in fact that the individual impact is hard to measure without seeing it as part of a large population doing the same sacrifice. Even someone as principled as Aaron realized that “it was silly to despair because people weren’t becoming vegans; most people were never going to become vegans, or even vegetarians, but if they ate less and less meat, then little by little things for animals were going to get better”. As such, the compounded impact matters more than the size of the individual sacrifice, whose value comes more from the inspiration it can provide as an example. Yet there is a fundamental difference between choosing what rules to follow (deontological framework), and how far to pursue a value (utilitarian framework):

“But it’s not only in its embrace of impartiality between family and strangers that utilitarianism is radical: it’s even more radical in how much it demands. In some rules-based moral systems, after all, there is the possibility that as long as you follow the rules, once you have done your duty, you are free to use the rest of your time and resources as you like. You haven’t stolen anything or murdered anyone, you have honored your parents, you have helped old ladies across the road, and now you can go about your business assured that you are a decent human being. But utilitarianism claims that you should act so as to bring about the most well-being possible in the world. Taken literally, this means that every single thing you do, at every moment of your life, should be motivated by that goal. Which is to say that **there is no point at which you can say, I have done my duty, I have followed the rules, and I am now free to do what I want.**”

There is no good solution to the trade-off between personal gratification and duty. The result is that this choice is arbitrary, and for a do-gooder, wherever the limit is set, any second spent not working is more chickens being tortured, every candy apple bought is the preventable death of a child somewhere.

Then there are the big sacrifices: “once you push moral commitment beyond its usual borders, more difficult conflicts arise: morality begins to push against other things that we value deeply, such as certain kinds of beauty, and certain kinds of freedom, and humility, and openness, and family, and love. To confront these conflicts is to ask what matters most in life.” The real sacrifices presented in the book are usually about family: most do-gooders seem to be beyond fearing for their personal safety, but asking a loved one to risk their life is a different matter. When his father gets sick, Aaron wrestles with whether his commitment to his family should come before his commitment to animals; but he knows that if it were his mother instead, whom he actually loves, he wouldn’t even think about the choice. Peter Singer himself decided to spend a lot of money to care for his mother when she developed Alzheimer’s disease. Which leads to the question of whether it’s ethical to have children when you are a do-gooder. First, there is the question of spending a lot of money on someone who doesn’t exist yet, instead of on suffering that exists right now. Both Aaron and Jen, and Jeff and Julia decide that

there is no moral justification for having children; the first couple sticks to their principles, and the second chooses to do the human thing.

Then there's the question of whether you can impose that way of life on someone who didn't ask for it. But this is a strange question, considering that the children who are living in the misery you're trying to reduce never asked to be born there either. The Badeau family is a useful concrete example. Every child who is added to the family becomes a strain on those who are already a part of it. And every time the parents want to make a new addition to the family, the children are asked about it. Welcoming the new child very likely means improving the life of their new sibling, no matter how thin the family's finances are stretched. Rejecting them means denying vital help to a stranger, one who is very much like them (the Badeaus almost exclusively adopt children who are unlikely to be adopted otherwise, either because they're a group of many siblings, or they have some disease or handicap). And yet they can't adopt every child in the world.

(Fortunately there are also easy decisions, cases where the do-gooders' values are aligned: of course nuclear weapons are an existential danger, but to Dorothy, they are mostly a waste of money that could be used to feed starving people. Animal cruelty is wrong, but eating meat is also less healthy, and the factory workers are some of the most ill-treated and miserable. Giving lepers their dignity is a worthy goal, but they also provide a valuable workforce)

It seems that avoiding those ethical dilemmas is a major reason for the undermining of do-gooders. Some of the criticism is about the effectiveness of the effort; not necessarily in the sense of EA, but just having an effect at all, or not inadvertently causing worse problems. This idea comes from disciples of Adam Smith, who considered that in a capitalist society, virtue was an anachronism, and capital could be better shared by letting the market decide. If anything was virtuous, it was to follow one's instinct to buy "stupid, extravagant, frivolous things": pride and greed were the new virtue. Although centuries of capitalism have shown the limits of the system, it is true that misguided attempts to help can easily fail, particularly if they're focused on the feelings of the giver rather than on the effects of the charity.

One common flaw in charity is ignoring the political dimensions of a problem. Sometimes the individual solution is cheap enough to be scalable, as with the case of deworming medicine (and even then, problems of coordination can ruin your efforts. An example of this on a smaller scale happened to Kimberly, who, with the help of some of her parishioners prepared a big meal for a family shelter, and ended up with lots of uneaten food because the shelter had forgotten to put it on the schedule). Other necessary reforms can only come from a top-down action from the government, or more ambitiously from institutional change.

Acting upon the object level without considering the factors that affect the situation is another mistake. Typically, food sent to countries facing famine can be exchanged for weapons on the black market, thus prolonging conflicts. MacFarquhar mentions Michael Maren's book *The Road to Hell*, which shows how food aid can destroy local markets, as well as Alex de Waal's *Famine Crimes*, that denounces the lack of accountability of

NGOs and how aid removes incentives for governments to prevent famines. In general, aid should empower the locals and give them the tools to develop their own solutions to the challenges they face, rather than make them dependent on neocolonialist programs. But all of this is more support for EA rather than an indictment of aid in general.

The problem that do-gooders pose is that they show that facing those hard dilemmas in a way that is courageous, thoughtful and constructive is possible: we “know, as the do-gooder knows, that there is always, somewhere, a need for help”. Do-gooders are a constant reminder of our own selfishness, an implicit criticism of how we should live our lives. The only way we can continue to see ourselves as good people under these circumstances is to discredit them.

One way to do so is to question their altruism: people are selfish, do-gooders are people, so they must be gaining something from their charity. Are they trying to feel like they are better than others? Do they want to feel powerful as they change the world? Are they secretly benefiting from the charity in some manner? Are they trying to assert their power on others by making them feel indebted and in no way to repay their debt? Of course with the rise of psychotherapy, such ideas took shape as a theory of altruism as a mental disease: “Selflessness was, in Freud’s view, usually suspect. The devoted, self-sacrificing mother, for instance, he found to be part masochist, part tyrant, enslaving her child with chains of guilt. But the devoted, self-sacrificing child was equally dubious.” This kind of reasoning led surgeons to be suspicious of kidney donations even in the case of a mother wanting to save her child: “Billy Watson (a pseudonym), a ten-year-old boy, needed a kidney transplant in order to live, and his mother wanted to donate. But was Mrs. Watson’s motivation acceptable or pathological? the doctors wondered. Mrs. Watson had nine other children—was she showing an unhealthy favoritism toward Billy by wanting to keep him alive, since the surgery would leave her temporarily unable to care properly for the others? And was this a normal family, psychologically? How stable was the Watson marriage?”. Or in another caricature of Freudianism: “after a sister donated to her brother, the sister “felt absolute control over her brother, as if she had castrated him.” Basically, any desire to do something for someone else can be framed as pathological: “narcissism turns morality into an autoerotic pleasure in which the pleasure itself will be suppressed.”

All this converges to the notion of codependency. The concept is developed in a chapter on alcohol. The relevance of this chapter is that according to MacFarquhar, “the recent history of do-gooders, at least in the West, is inseparable from the history of alcohol”. This probably explains why along the relationship of every character to religion, she made sure to also mention their relationship to alcohol, whether it is becoming an alcoholic, having grown around alcoholics, or categorically refusing to touch alcohol, like Jeff, Aaron and Baba’s family (nothing culturally forbids them, except for Baba’s edict on the matter). She writes that when you drink, “you feel more empathy for your fellow, but at the same time, because you are drunk, you render yourself unable to help him; so, **to drink is to say, I am a sinner, I have chosen not to help.**” I can’t decide whether it is her position or if she’s trying to express the perspective of an imaginary do-gooder, but in that case, all we have is Jeff and Aaron’s perspective on the topic, who never drink for fear of the loss of control and the alteration of their mind; while I understand how that



makes sense as an argument against being drunk, I'm not sure why they couldn't have one drink. But then again, do-gooders rarely compromise unless they have to.

More precisely, this chapter is about Alcoholics Anonymous (AA). In 1934, Bill Wilson had a religious vision that prompted him to stop drinking and start AA. Throughout his alcoholic phase, his wife Lois endured everything and tried to make him stop, in vain. After he overcame his disease on his own and remained sober through the support of other drunks and recovering alcoholics rather than through the power of her love, she starts thinking she might actually have been the problem. When Bill was drunk, he was dependent on her, now, she had lost her purpose. She made this codependency model the centre of her own organisation, Al-Anon, a support group for the families of alcoholics. Al-Anon teaches its members that the actual problem is their self-righteousness, they are "co-alcoholics"; they need to acknowledge their problem and follow their own twelve step program.

Here MacFarquhar bizarrely jumps from this individual model of codependency from Al-Anon to making it the template for the criticism against do-gooders, regardless of the problem they're trying to solve. Is she saying that people think those who fund MIRI are codependent with the AI alignment problem? That Aaron is trying to gain power over chickens? The examples presented in the rest of the chapter are not this extreme, and it is conceivable that some of the people who do disaster relief abroad might have some savior complex for instance, but as a general theory of the hate against do-gooders, this chapter falls short.

One last reproach against do-gooders is that if we are to take them at face value, without supposing hidden or unconscious gain from their endeavours, this makes them truly inhuman beings. For instance the chapter on voluntary kidney donation mentions how "bizarre, even repellent" the idea appears to the surgeons: doesn't it go against the Hippocratic oath? A study on the phenomenon concludes that "the most puzzling aspect of our investigation has been the striking contrast between the naturalness, relative calm and equanimity of the volunteer donor, and the uncomfortableness of the transplant team."

When belief in God and Heaven was a fact of life, someone with the mentality of a do-gooder could be seen as a hero, someone to emulate: if doing good is the path to heaven, then giving up a bit of comfort to avoid hell was only sensible, and we can only wish we were strong enough to do the same. But in our more secular societies, doing good to such extent, just for the sake of living a moral life, without expecting any reward, even in the afterlife... it's strange. Reviewing Gandhi's memoir, George Orwell wrote that "the ordinary man only rejects [saintliness] because it is too difficult [...] Many people genuinely do not wish to be saints."

Normal humans are morally imperfect. Although Singer's drowning child thought experiment is easy to understand, we cannot follow it through our actions. "Surely, a person who declined to reduce himself and his family to penury could not be as bad as a serial murderer", yet it's difficult to express the theoretical difference. Except that accepting that fact when we are incapable of moral perfection would mean renouncing



the concept of morality altogether: “to demand too much can be as corrupting as demanding too little. And to demand too much crushes aspiration. If every good act is required, then none is praiseworthy. There is no more virtue—only duty and vice.” Then the only moral path is to devote our life to an “impartial, universal love” (with little in terms of transcendent joy), which “seems the antithesis of what we value about deep human attachment.” There is thus no moral life for a human being. We don’t need to adopt such an extreme view: simply seeing the example of the imperfect do-gooder (who still loves their family more than the stranger) and then choosing not to follow it means that we have chosen not to be as ethical as we could be.

And even if the do-gooder still values their family more than the stranger, the cold rationality that makes it possible for them to even compare the two is already repellent to some:

“Suppose a man could save either his wife or two strangers from drowning, [philosopher Bernard Williams] proposed. A utilitarian might ask: is it permissible for the man to save his wife? But even to ask that question in such a situation was to have, as Williams dryly put it, “one thought too many.” A loving husband would save his wife spontaneously, without consulting moral rules at all. To grant morality the power to adjudicate impartially in situations like that would be to abandon what gives human life its meaning. Without selfish partiality—to people you are deeply attached to, your family and friends, to place—we are nothing. We are creatures of kinship and loyalty, not blind servants of the world.”

Even without introducing family into the equation, the fact that there could be an equation involving human lives makes people uncomfortable, even when it’s trading money for human lives rather than the opposite. MacFarquhar postulates that after the French Terror and the Russian Communist Revolution, the idea of a small group of people convinced that they know the path to a better world and willing to sacrifice themselves for it can only make us think of death, and of whether we would even have a place in such a world.

### III.

One important thing to realise about the do-gooders presented in the book is that their sacrifices were not empty. Large sums were sent to such charities as the Against Malaria Foundation, corresponding to hundreds of lives saved by the people mentioned in the book. Many more survived and regained some of their dignity in Nicaragua, India or Senegal through direct interventions. Some suicidal people regained some will to live. Legislative and corporate regulations were decided to improve the well-being of chickens. And although the story of the Badeaus is not presented as the same unmitigated success, with three of the children predictably dying of their respective diseases and two ending up in jail, all of them had a better life than they would likely have had they not been adopted. Even Charles Gray might have come close to partly achieving his goals, as the fast he and his friends conducted in Paris, Bonn and Oakland attracted political attention in France and Germany. Had the Cold War not

reheated just at that time, some small progress toward nuclear disarmament might have been achieved.

In the last chapter, MacFarquhar reminds us that there were always limits to the sacrifices the do-gooders were willing to make, and that despite how unwavering they appear in their commitment, they still doubted themselves (which she sees as “a measure of their seriousness”). And although they may appear to us as extraordinary, it’s only because we see them through our own norms. These norms certainly made their actions more difficult: it would be easier to be generous if such behaviour generated praise rather than confusion, suspicion and hostility. Luckily, such norms do evolve: the chapter on kidney donations points out that this act that used to be the clearest possible sign of a diseased mind has now become, if not normal, at least acceptable enough to have processes in place to accept those donations (still, making it an opt-out process rather than opt-in would be such a simple way to save lives, it’s really frustrating that all countries don’t do so). Similarly, as EA becomes more prominent, we can imagine a world in which nobody would take the [Giving What We Can pledge](#) because there wouldn’t be anything special to giving 10% of one’s income past a certain level of wealth.

But what is normal is relative. Kimberly’s chapter is a good example of that: she was an aid worker in dangerous countries? How is that different from what Dorothy and Baba did? She adopted two children, one of whom was handicapped? That’s only a tenth of what the Badeaus did! She gave a kidney to a stranger? Thousands of people have done that! Have I achieved anything close to any of that? Not at all. But simply the fact of being presented near the rest shifted my expectations of how a moral person should behave. Similarly, if you agree that you should give more but find it difficult to commit to that decision, I recommend spending time among EA circles and reading EA blogs and books, until peer pressure pushes you to do the right thing. Then keep looking up to more extreme examples until you reach the point where you’re no longer comfortable giving more. Or you might become like Aaron who apparently ended up finding it hard to meet more inspiring people:

“He still needed to accumulate a certain number of utility points each day to feel okay with himself—to feel that he’d done his duty—but since his work was going so well, those points were easier to come by. It had never occurred to him when he was younger that there might be a time in his life when he might feel he was doing enough. Mostly this was a good feeling, but he was also suspicious of it.”

But can anyone be a do-gooder? Or is it essential to be a religious teetotaler from a hard background, as MacFarquhar seems to suggest? The last bio in the book is that of Stephanie, who after working for [GiveWell](#), ends up rejecting effective altruism, as she couldn’t stand the constant guilt of not doing more, or the fact that more personal but less optimal ways to do good were discouraged by the movement. She ends up finding peace through self-acceptance. She decides that the fact that logically she should be trying to save every drowning child, doesn’t mean she has to: she can also find her own ways to do good. Although do-gooders tend to live by some rules, these are self-

imposed, and it's OK to change them if you can't find happiness through them: getting embittered would help no one.

Similarly, even though Julia used to think that other things than her own happiness mattered, she ended up deciding to have a child and keep her job; it doesn't have the highest impact, but it does good and makes her feel good about herself. She relays Eliezer Yudkowsky's advice to [purchase fuzzies and utilons separately](#): that doesn't mean you shouldn't purchase fuzzies at all! Even from a utilitarian perspective, being a happy altruist is more likely to convince other people to give than being an effective but miserable one. The people you can encourage to give are part of your impact, and it is thus useful for the EA movement to be as inviting as possible, typically by being welcoming of other people and encouraging people to be generous, rather than judging them for not being the most effective. The goal of charity is to improve the world, and this is more easily achieved by having more people contributing: we wouldn't have to witness all these strangers drowning if more people were helping. As Julia says:

"One thing I almost never talk about is anger. When I'm happy with my life I don't have any reason to feel angry, but when I'm feeling deprived I sometimes do. I feel like I'm pulling at something heavy that I can't possibly lift by myself, maybe pulling a car out of a mud pit. And everyone is standing around saying, "Boy, it's too bad that car is in that mud pit," or "That looks like hard work you're doing," or, more often, "Did you hear the Italian team just lost out to Slovakia?" I really think there's enough material stuff and human ingenuity that nobody needs to be horribly poor. If everyone who could, pulled a little more weight, I wouldn't need to pull so much."

So thank you for doing good, in whatever way feels right to you!

# Strong Towns by Charles Marohn

For years I've wanted to open a restaurant. The plan goes like this: Every Saturday morning I'll put a folding table up in my front yard and sell vegetables from my garden and eggs from my chickens. Once people are in the habit of coming by, I'll add some baked goods. If those catch on, I'll throw up a picnic table and serve breakfast. And I'll keep going until I have a small, part time restaurant on my lawn. Each step will be small and low risk. If people don't come by, I'll stop and move on to something else. It's a risk free way to pursue my dream.

Unfortunately, I'd be dragged away to an HOA gulag on day one.

In Strong Towns, Charles Marohn points out that this sort of incremental progress used to be the driving force behind all city development. Things we would identify as 'businesses' like hotels, restaurants and stores would grow out of someone's house. If the business failed the owner would shrug and turn the space back into a living room. Or rent it out. Or turn it into a workshop. Because *no one cared what you did with your house*.

Modern America does things slightly differently. We build stores, malls, and entire neighborhoods to their finished state without worrying too much about the outcome. If a big box store fails, we just close it. If no one moves into a development, the developer slashes prices and takes a loss. The city, who subsidizes these places and is on the hook for maintaining them, loses some money, but that doesn't seem like a huge deal. Marohn points out here that as a culture, we're not used to worrying about resources; we've had too much [slack](#) for too long. The book doesn't provide much to back up this assertion, but "we've had too much slack for too long" feels like it could be our national motto right now, so I bought it.

Everything we *do* build gets frozen in place. Not just with zoning regulations, but with environmental regulations, referendums, and every other tool of [vetocracy](#). Everything is either undeveloped or in its final state; there's no middle ground.

This may not be a universal intuition, but "take a complex adaptive system like a city and start removing degrees of freedom," seems like it might end badly.

---

Marohn spends the first forty pages convincing the reader that the American system of urban development is a departure from historical norms. He shows off several examples of how incremental growth occurred in the past, then explains that automobiles allowed us to change the fundamental structure of cities by allowing people who work in the city center to live on the outskirts, devaluing the neighborhoods near the core. I found myself nodding along. This at

least matches my personal observations. The most car friendly cities I've lived in have a ring of decaying neighborhoods that suburban commuters speed by on the way to the city core.

After making this case, the book goes straight into telling the reader that most American cities are heading towards bankruptcy. I'm not exaggerating. The most apocalyptic quote was:

"Detroit is not some strange anomaly. It's just early. It's just a couple of decades ahead of everyplace else."

Marohn's case is simple: a city needs to take in as much money from property taxes as it spends on maintenance, and not only are US cities not meeting this bar, they're not even trying. Not because the people running them are corrupt or incompetent, but because 'making money' isn't an intuitive thing to optimize for if you're running a city. After all, isn't a city government's primary goal to make sure the citizens are happy and taken care of? It'd be career suicide to say "we're not going to repave your roads because your neighbourhood is a huge financial drain on the rest of the city." But if your city is operating at a loss you *need* to write off certain neighborhoods (or increase taxes, which is an equally popular way of making ends meet).

(This section reminded me of the culture at Google. If you want to get promoted you work on one of the big flashy projects. No one is going around saying "we could make a new chat service, but what would be *really* great is to allocate those person-hours to improving our core services.")

Every time the government does something like add a lane to a highway, they have added a permanent, ongoing expense to the city. This means that, in order to be sustainable, improvements to the highway need to either increase property values enough that the road pays for itself or city managers need to raise taxes somewhere. Even a one time cost like resurfacing a road or cleaning up garbage has to cause a corresponding bump in revenue or that city will never get rid of the deficit.

This insight floored me. It never occurred to me to think about where the money from city projects came from. Building roads and parks is just what the city did, like some benevolent fairy reaching down to make sure the lives of rich white folks ran smoothly. Thinking back on the construction projects I've seen, they range from improving home values marginally ([we tore down this building and replaced it with a field of astroturf!](#)) to actively hurting city income (we tore down hundreds of houses to run a noisy highway across the county so people can get to work five minutes faster!). Thinking about it this way, I started to wonder how our cities have lasted as long as they have.

Marohn's answer to this question is two-fold. The first one is pretty obvious: cities are taking on unsustainable levels of debt. He explains why city management might make this mistake, why lenders might still finance it, and why debt is bad in certain places and not in others. This section felt unnecessary; I'd be more surprised if cities *weren't* taking out unsustainable levels of

debt. As it was, I read this section thinking "yeah, that's what people do when they don't have enough money."

The second answer is more interesting: Cities are financing themselves through unsustainable growth. Because maintenance costs don't really become significant for 20 or 30 years, a city can increase its tax base by building new developments without paying any immediate cost. But as the pipes and roads age, the maintenance costs of those developments go up while the property values go down. Why wouldn't they? We keep our neighborhoods static, so there are limited ways people can improve the values of their property. And this same unsustainable growth increases the supply of housing, further preventing home prices from rising. Thus, every development the city builds eventually becomes a liability, which prompts more growth, etc.

(I would have loved it if he had talked about the developer perspective. It seems like a good way to make money would be to create built-to-completion neighborhoods that look as impressive as possible while saving money by not building them to last. I haven't really heard of anyone actually checking up on the reputation of the people who built their house, and besides, if it takes ten years for your houses to fall apart you can probably sell a heck of a lot of them before anyone catches on. I don't know if anyone is doing this persay, but I'd be surprised if they weren't.)

The fact that city planners seem to value the least cost efficient properties further exacerbates the unsustainable growth problem. Franchises and big box stores seem flashy, but are, according to Marohn, extremely inefficient when you consider their infrastructure use relative to their tax revenue. Assuming acreage is a decent proxy for maintenance costs (which I'm not sure I fully understand, but I'll trust him here) he gives us the tax revenue per acre of various city sites. Big box stores and chains come in last; areas that are considered 'blighted' and tend to be targeted for redevelopment are some of the highest performing properties.

All this taken together paints a pretty grim picture for our future. If Marohn is correct, cities have been continuously adding on liability for decades and only staving it off with constant growth. At some point, the accumulated maintenance cost will outpace a city's ability to grow and the city will be forced to either face bankruptcy or contract, abandoning some neighborhoods to neglect.

So all our cities are going bankrupt and we're going to live in a decaying suburban hellscape. What should we do? Marohn has some suggestions.

Allow cities greater autonomy to govern themselves. Stop building new infrastructure. Identify what the highest earning parts of the city are, focus on them, and commit to abandoning low performance areas (which, to him, means areas that produce the least tax revenue relative to maintenance expenses). Limit the debt cities can take on. Revitalize old or vacant areas by letting people use them however they want. And get rid of regulations and zoning rules that aren't absolutely essential.

These suggestions encompass the later quarter of the book and are clearly well thought out, but I'm not going to go too in depth with them for two reasons: One, I think a lot of the solutions are on a "once you realize you're in a hole, stop digging" level. And two, if you're in a position to actually implement any of the changes he suggests, you should probably read the book yourself. That being said, if you're one of those few who are *not* in a position to say, pass a state law limiting how much debt cities can take on, there are a number of non-government organizations trying to improve cities at the community level. There is a tactical urbanism movement that focuses around small scale interventions and has a [Wikipedia page that is mostly a list of grass root projects that can improve the city](#). There is also the [Better Block Foundation](#) that focuses on improving urban spaces, and the [Strong Towns](#) movement that focuses on building prosperity (to Marohn's credit he does spend the book beating you over the head with how awesome his org is.)

The beauty of Marohn's suggestions is that they continue to be applicable even once his predictions start becoming true. Unlike global warming, where once you can see the most damaging effects it's already too late, a lot of what he describes here can be employed at any point during urban decay. While I think addressing these issues before they become glaring problems is unlikely (because why would we ever address a problem before we start experiencing the consequences), it's nice to know that when this *does* start becoming obvious to people I'll have a book to lend them.

---

The last section of this book shifts in tone from practical to Marohn's feelings about the actual quality of life in a suburb vs an urban environment. He argues that by abandoning typical development patterns we've also warped our social structure, that suburbs are just not conducive to a healthy tight-knit community. I'm sure it depends on the suburb, people and the culture, but I don't think he's wrong that something in how we relate to our neighbors has changed.

Towards the end Marohn describes his experience moving from a suburb to a healthy urban area. One passage in particular resonated with me so deeply I wanted to include it in its entirety:

"At the old house, the girls were stuck biking the driveway, maybe the cul-de-sac. Anything unaccompanied beyond that risked tragedy with a speeding vehicle and a moment of inattention. And even if they had ventured further, there was no place to go. Their friends were likewise spread across the area, brought together outside of school only through scheduled play dates and long car rides.

The new neighborhood is full of kids. A city park is a mere block away. The girls have learned to navigate the neighborhood on their own, as a life skill I took for granted yet few of their peers seem to possess. They can go to the store or the library or, when they scrounge up some

money, downtown for an icecream. I'm watching them grow into the kind of confident people I always hoped they would be.

The move has saved me five hours a week of commuting. Instead of being stuck behind the wheel, I walk or bike to work. And to get groceries. And to the hardware store. And anyplace else I can reasonably get to, even in the winter. I go days without driving, and I've felt myself become more relaxed, less stressed, as a result.

The most consequential change with the move has been adding new neighbors to my life."

He continues:

"I'm an introvert who enjoys long walks alone, avoids social gatherings where possible, and can't remember names. None of that has kept me from getting to know all my immediate neighbors. It's almost impossible not to as we run into them all the time. I know their kids, their pets, and some of their plants. When we're out of town, they watch our house. We do likewise to them."

I grew up in a suburb that is somewhere in between suburban sprawl and Marohn's ideal. We knew a few kids in the neighborhood and could, if you were tenacious, walk to the store. But even so, there's something about suburbs I find isolating. The houses are their own islands, separated by space and fences and trees. It doesn't feel like an environment designed to foster human connection, but rather to squash it.

The life Marohn describes seems idyllic enough to me that I feel almost cheated by my lifelong existence in the suburbs. Reading this book, I have to wonder if some of the social decay we're seeing now is the result of the transition from living in tight knit communities to living in our own fortress-like cookie-cutter houses, surrounded by a green grass moat trimmed exactly to HOA specifications.



# Stubborn Attachments, by Tyler Cowen

Growth is good. That was the first and was sufficient for the last statement of the book. Unwilling to let perfectly fine statements stand by themselves, the foreword goes on to explain what that means, and why they think that is true. Not just growth in the general sense, but economic growth specifically. Not just good in the general sense, but ethically good, from multiple philosophical frameworks. Not just 'is' in a general sense, but 'is' in a way that connects words!

Tyler Cowen himself is an excellent economics professor. He also authors the blog *Marginal Revolution*, which is a very good place to find sane analysis of important events in the world, along with links to other well written pieces. His podcast, *Conversations with Tyler*, is worth the listen, and branches from Cowen's usual focus on economics to culture as well.

On the book itself, the subtitle gives more clues to the subject of the book, certainly more so than the title. "A vision for a society of free, prosperous, and responsible individuals" sounds like the sequel to *Atlas Shrugged* at worst, but a very nice place to live at best. The nice part about this book is, purely on the author's credentials, I doubt that I am going to hear the rant of a person who read a book on philosophy. That benefit of the doubt is important, as the subject of what society should do gets tackled by almost everyone at some point, so we are dealing with a very crowded field.

Tyler comes at the idea from a rather libertarian perspective. His core values for the book are prosperity and individual liberty, which do not sound so bad at all. Who doesn't want wealth, and the ability to do as you please? He claims the existence of an objective right and wrong, rather than a relative right and relative wrong. He also doses us with a reminder to not devolve into "us vs them" thinking, remain skeptical of panaceas and other forms of common logical fallacies which are so common in the minefield of societal goals. Most entertaining, he sets up thought experiments where the solution obviously presents itself, in a way that direct argumentation does not. We need to read these philosophical underpinnings because Tyler then presents controversial society-wide problems, along with his proposed solutions which should inspire us to act differently, if all goes well.

## In Search of A Free Lunch

Tyler introduces the concept of a mythical *Crusonia* plant, which is a crop that creates more and more each harvest. The point here is that decisions with tradeoffs are hard,

but we should prefer choices which give us more and more over time instead of choices which give us one pay off. Everyday examples of this kind of choice are hard, because nothing is a true perpetual motion machine. However, there is a clear tradeoff between eating your crop seeds and planting your crops so you get more food later. The idea being, you get more of a free lunch by choosing an option which has a greater payoff in total rather than picking the soonest possible payoff. Education could qualify here, assuming the education that you get is actually useful to you in the future. One very hard part of figuring out which decisions will have very long term payoffs is that knowing for sure what the future holds is very hard. I think that is a good reason why picking the short term payoff is usually easier, and seems to be picked more often. The hardest example here is Chernobyl. Should we occasionally have a Chernobyl event so safety in nuclear power plants becomes more important? The short term payoff of not doing that is positive, in that we continue to produce energy, but if we prevent two other meltdowns from happening, then Chernobyl was actually a good thing, and ultimately leads to more energy production later.

The difference between short term and long term payoffs becomes obvious in sustainable compared to non-sustainable economic growth. The Soviet Union famously had growth rates of 10% or more in the 1950s, but they did so by consuming larger and larger quantities of raw materials as inputs to their factories. This becomes a problem if you run out of resources, and find that your economy collapses shortly after, resulting in misery and poverty (Not that U.S. could tell at the time. The 1950s were intense!).

Therefore, we cannot only pursue breakneck economic growth at the cost of sustained growth, and indeed not even at the cost of all leisure time. This is an extremely difficult tradeoff. Certain people seem to enjoy working 16 hour days in perpetuity, while others become miserable doing the same. Our growth rate would certainly be higher if everyone was working during all of their waking hours, but we might eventually touch on the nebulous concept of burnout. We are again left trying to predict the future results of current actions, looking at the tradeoff between working extremely hard and risking burnout, and taking more leisure time but potentially missing out on higher sustainable economic growth. Personally, I don't think I have ever experienced burnout. At worst, I remember working a full time job, and taking calculus, proofs, and data structures classes. I don't think I could have worked any harder than that without sleeping less, but I don't remember hating the idea of working even at its worst. I believe other people when they say that they are burned out, and I do not think pushing them to keep working would necessarily be good. Knowing this point is important because the long term growth rate is partially composed of human labor.

Despite the uncertainty on how to grow Crusonia plants, Tyler gives excellent examples of why we should strive for growth. The past was really terrible and not at all idyllic,

despite how we might romanticize it. Tyler walks through many interesting examples of precisely how bad things really were, but I think the most moving was “...life expectancy in Western Europe was roughly forty years of age, and food took up fifty to seventy-five percent of a typical family budget. The typical diet in eighteenth-century France had about the same energy value as that of Rwanda in 1965” (34). This is a remarkable consequence of growth. Would you rather have been born in France or Rwanda? It is our moral failing that we can produce arbitrarily large amounts of food today, and people still starve, but our world would be better still if Rwanda had simply been growing at a high enough rate to produce cheap food for itself.

Finally, Tyler touches on the most cursed area of science: happiness research. I call it cursed because happiness seems insanely illusive, and many things that we immensely value have very little benefit to happiness. Tyler cites many studies in support of his point that growth could positively affect happiness due to knock on effects of being less likely to be crippled by disease or accident, and envy is not that bad, given that you can still enjoy better products in a high growth society. The problem is that for each of his studies that he cites, another study can be cited showing that poorer countries are happier than richer countries. I think that going to happiness research is a mistake to prove a point, because happiness never seems to change enough to justify action. As an aside, I could not find the origin of the term Crusonia. Tyler got the phrase from Frank Knight, a 20th century economist. The problem in the book is that Tyler says that the idea comes from a novel called Robinson Crusoe Island, published in 1954. However, this citation dates back to 1944, so somebody has their dates mixed up. Until I read the original Robinson Crusoe Island novel, we can safely assume that the Crusonia plant comes from Harry Potter.

## Comparing Desires for Apples And Oranges

How do we compare different desires? If two people want an apple, but you only have one, how do you pick which person becomes better off if they get an apple? This is a classic aggregation problem, and gets at another protest to economic growth. How do we truly know which option to take, if the tradeoffs on different people's preferences cannot actually be measured? Tyler takes the approach that choosing the option which is more likely to result in long term economic growth is far better. Choosing higher economic growth will result in the tradeoff disappearing, as everyone can have more of what they want. In the apple example, suppose you chose to plant an apple tree instead of giving the apple to either person. In a few years, both people would be able to enjoy far more apples than the original tradeoff implied, so Tyler thinks we should prefer the growth option, and satisfy everyone's preferences later rather than take a tradeoff

today.

I think this can often be easier when the games that we play are not necessarily zero sum. We cheated in the apple example by making more apples, and economic growth allows us to cheat competitive games in the same way. Perhaps the most important point here is to move towards situations where you have the option to pick growth, and away from games with finite resources

Economic growth seems to allow us to escape poverty, misery, and even bad competition, but Tyler introduces the constraint of 'Don't violate human rights'. He is mostly talking about negative liberties, or the idea that you cannot stab, torture, kill or otherwise abuse another person. Imagining an example where torture increases the long term growth rate is difficult, but stealing or even compelling people to do things is much easier. Microsoft has a fairly high rate of return, while poor people have low rates of return on their money. Should we therefore steal from the poor and give the money to Microsoft, so that we get a higher rate of return on money? (probably not).

I am all for respecting human rights. Killing and stealing for the sake of economic growth seems really bad, even if the benefits of long term growth would outweigh the costs. There is still something repugnant about harming someone in order to achieve a goal, and you immediately are back to a zero sum game of stating that the overall pleasure of other people is worth harming someone. Since solving these kinds of harm/benefit problems does not seem to be possible, we would need to avoid doing that in the first place. I am still interested in seeing an example where killing someone actually improved the growth rate. No real events come to mind, and judging the benefits of killing someone is gruesome work.

## Time, and Other Illusions

Tyler thinks we should value the future just as much as the present. In one of the weirder thought experiments on the internet, he argues about relativistic astronauts. If we built a rocket that was going to approach the speed of light, the astronauts in time would have their sense of time slowed relative to ours. A normal trip in their time would be eons in ours. Should we build landing gear on that rocket ship? After all, we will never experience any sort of benefit or harm from knowing that a rocket will explode millions of years in the future. In order to answer yes, which seems to be more sane than no, we would have to have no discount rate on the value of future lives over our own. Apart from space ships, Tyler gives more grounded examples as well. On average, regions which were prosperous even as far as 2,000 years ago are more likely to be prosperous today. This presents a problem. Clearly, we cannot go back in time to pay our ancestors to work harder and make good choices, but we very much benefit

from them having done so (or not done so).

In particular, this situation reminds me of the Cuban Missile Crisis, and other borderline catastrophes, during the Cold War. A hot war would have reduced civilization to rubble in a few days, killing most or all humans on Earth. We would be willing to pay a lot of money, or do anything else, in order to prevent a hot war, but we never got a vote in the matter because we were either not alive, or not one of the people who were deciding whether or not to start nuking. By allowing tensions to get so high, the USA and USSR were discounting the lives of those reading this very heavily, and if they had actually fired, that discount would have gone to 0. It is hard to make the same case for the future, because we do not know what is actually happening in the minds of our world leaders, and whether or not they are planning a war.

Apart from war, we enjoy a very high standard of living thanks to growth which started and mostly happened before we were born. We are much more able to affect the economic growth rate, however modestly, so working on that will provide a massive benefit to our descendants hundreds of years in the future. If we feel at all altruistic towards the future, we should therefore treat future lives with the same importance as present lives. Finally, it is possible that there will be many more people in the future than there are today, and many more to benefit from higher economic growth. We start to get into topics of existential risk, which would make all of this pointless, so assuming that nothing really bad happens, growth would be one of the most ethical things we can do.

## The Common Sense Approach to Selling Your Baby

Philosophy can get weird. The problem with most ethical frameworks is that, somewhere in there, you eventually happen upon an idea which most people use. We do not follow that idea to its logical conclusion because that would be insane, although pinning down why is hard. For example, there are many starving babies which you could feed for not too much money. If you sold your baby, then you would be able to save many other babies, and the world is a better place...or something. Finding that logic appealing might mean you would really benefit from psychiatric help, or you are really into philosophy.

Tyler claims that his law of maximizing growth subsumes the need to give money to the poor. If helping the poor today will increase growth, then this would already be what we are doing. If doing so would not help with growth, then then we should not take that action. This can make us blanch. The problem with such an action is that helping people in the most dire poverty will probably not mean that they will go on to found massive corporations and contribute to growth. Do we owe anything to the people who will die

young simply due to the circumstances in which they were born? This is a very difficult question to answer, especially if we have an obligation to future generations, whose suffering we cannot clearly see. Tyler does believe in adequate nutrition and education on the grounds that this will make society more stable and potentially give knowledge to people who will go on to cause economic growth. This mitigates the absolute tradeoff between today and tomorrow, but still causes issues for people who will most likely not be productive in the traditional economic sense.

This chapter is very much worth reading, as we implicitly make tradeoffs in the wellbeing of others every time we spend money. It helps to be more explicit in specifically what we should be doing, and why because then we will be more clear on what our own values are, and which sacrifices we are willing to make.

## Butterflies and Hitler

As a result of very small actions, our future is drastically changing, moment by moment. This makes acting on the basis of economic growth and concern for the future immensely difficult. The Butterfly Effect made this idea famous, with the idea being that prediction in weather is so hard that a butterfly flapping its wings could eventually cause a hurricane. Likewise, if any one of Hitler's ancestors had died, or made slightly different actions, Hitler never would have been born and we would not have seen the horrors of World War II and the Holocaust. Therefore, does anything matter. Things probably matter, and Tyler argues this by pointing out that we rarely need to consider what he calls the froth of uncertainty. Since all of our actions are likely to spawn butterfly effects, no matter what we do, as long as we respect human rights and attempt to increase growth, the large effects of our actions will on net drown out the smaller effects. This is always hard to measure. How do you know that you are not currently causing something terrible to happen by accident? Since the future is uncertain, and we cannot usually check the future against the counterfactual exactly, literally measuring the results of our actions is impossible. I think that this is a very vague argument, which Tyler seems to be addressing only to stop a very nihilist critique of his work. We cannot stop affecting the world, and even doing nothing is a choice, so uncertainty over the future seems to be a pointless critique.

## Finishing the Philosophical Conversation

So where does this book lead us? Tyler wants us to try for larger economic growth over

very long time horizons, with the only caveat being that human rights are inviolate. What I always find confusing at the end is “So what?”. Tyler does an excellent job arguing for his case. If he is completely convincing, then we have ourselves a new political philosophy, one which values the future, charity, growth, and autonomy. We do not know with high certainty how to bring that philosophy into the real world, nor do we know exactly how this philosophy should affect many of our personal actions. What we do have is a very hopeful future, which we can work on bringing about in small ways, having had our common sense adjusted towards doing more good in a broader sense. I think throwing in with Tyler here is wise, as we do not know just how good life can get. We do know how bad life was, so our ultimate act of generosity would be to seek new ways of making sure the future is much better than the present.





# Ten Days That Shook the World by John Reed

I

My copy of [\*Ten Days That Shook the World\*](#) starts with an introduction by A. J. P. Taylor, a celebrated British historian and socialist. He explains that the British Communist Party rejected the introduction when he first wrote it-- it was published only after the Party's copyright expired. He goes on to say that John Reed wrote *Ten Days* in 1918 in New York based on notes from Petrograd during the revolution of October 1917 (November in the modern calendar); that Reed was a "passionate Socialist" and outspoken supporter of the Bolshevik revolutionaries; and that Reed invented many details based on what his Bolshevik sources told him. Barely three pages in, Taylor has already warned his readers not to trust most of the book they're about to read.

I'm grateful for his warning. *Ten Days* isn't a dishonest book, but the way Reed wrote it makes it unreliable at best and misleading at worst. What it really cries out for is context. Large parts of it didn't even make sense to me-- let alone win my trust-- on the first reading. So, following Taylor's example, I did more research. I read [\*A People's Tragedy: The Russian Revolution 1891-1924\*](#), a one-volume history by Orlando Figes. Then I went back, read *Ten Days* again, and wrote this review.

In the end, I feel like I've read *A People's Tragedy* so you don't have to read *Ten Days*. This is still going to be a review of *Ten Days*, mostly, though I'll be relying on what I learned from reading *A People's Tragedy*. (Figes has blind spots and biases of his own-- roughly the ones you'd expect from a Western academic writing in 1995-- but I'm using him mostly as a source for the historical consensus. In some places I've also interpreted or simplified his conclusions.) What I realized while reading *Ten Days* for the second time is that Reed's work doesn't add much to what Figes presents. It's a masterpiece of long-form journalism or propaganda or both, depending on who you ask, but as a source of historical understanding it's unhelpful. The interesting part of reviewing it is asking what Reed thought the story was, how he came to see it that way, and whether his framing is a useful one.

II

*Ten Days That Shook the World* starts about a month before the days in question, which isn't nearly enough. To make sense of what's happening we have to go back to, say, the early 1900s when the Tsars were still firmly in power.

Let's be clear: Tsarist Russia in 1900 was a miserable place to live. People like to joke about the Soviet regime-- *In Soviet Russia, party can always find YOU!*-- because Soviet awfulness was funny. It was novel and ironic and not what anyone expected. Tsarist Russia wasn't novel or ironic. It was awful in exactly the ways everyone would expect, because it had barely changed for the last three hundred years. In a rapidly-modernizing world Russia was stuck under an honest-to-goodness feudal divine-right autocracy. Just imagine the punch-lines:

*In Tsarist Russia, alcoholism and domestic abuse are fixtures of village life!*

*In Tsarist Russia, religious leaders are mostly enforcers and propagandists for the monarchy!*

*In Tsarist Russia, the outdated military gets trounced by tiny, barely-industrialized Japan!*

*In Tsarist Russia, a corrupt illiterate mystic can control national policy!*

*In Tsarist Russia, troops routinely fire indiscriminately on unarmed protesters!*

Horrific, yes, but not so chuckle-inducing.

Here's something I hadn't realized before, though: unlike our usual picture of an oppressive government, the Tsarist state was weak. Its bureaucracy and communications and decision-making process weren't remotely up to governing the country that Russia was becoming. The traditional agents of the monarchy were the landed nobles, who were effective at managing big rural areas without a good communication network or an educated bureaucracy. Up until, say, the mid-19th century that's exactly what Russia was. But eventually Russia had to industrialize, because it turns out rival nations walk all over you if you don't. And industrialization created social pressures that the Tsars couldn't, or wouldn't, address.

Those pressures were tied to the growth of cities, which had an educated intellectual class and an industrial working class. The Tsars hadn't really had to deal with those classes before, and they did a terrible job now. At first they tried to ignore them, then they tried brutal repression, and finally-- in 1905-- they half-heartedly tried to liberalize by establishing a Duma (parliament), which in practice they ignored. The trouble, as Figes tells it, was that the last Tsar-- Nicholas II, who reigned from 1894 to 1917-- never really understood the seriousness of the situation. He'd bought into his own myth of divine right, and his instinct was to double down on autocracy when push came to shove [p. 275].

What you end up with, by 1915 or so, is a Russia that still hasn't industrialized much, let alone come to grips with the social changes industrialization brings. The urban lower class lives in terrible conditions, including frequent food shortages. The urban upper class wants reform badly, but is constantly stymied by the Tsars. The cities are full of intellectual fads imported from the more industrialized parts of Europe, including, increasingly, Marxism. The Duma has given all of them a taste of politics, but the liberal institutions common in Western Europe haven't had a chance to develop. The Tsar and his cronies still have absolute power, and they refuse to

engage with the urban classes, seeing them only as a threat to the monarchy. Their idea of a helpful response is to pick stupid fights with Japan (in 1904) and the Central Powers (in 1914) that are supposed to gin up patriotism, but only worsen Russia's economic and social problems.

Then there are the peasants.

The peasants are the elephant in the room in any discussion of Russian politics, even though they're usually not actually in the room. (Socialists from the lower classes tend to be ex-peasants who got out of their native village as soon as they could. Figes argues [p. 110] that they tended to "reject and even despise" peasant culture for ideological and class reasons.) They're 80% of Russia's population and mostly illiterate. They're also, by any reasonable standard, dirt poor. In fact even the dirt is kind of a question mark: serfdom only ended in 1861 and the nobility still owns a substantial fraction of the rural land.

Peasant society is, depending on how you squint at it, socialist, anarchist, traditionalist, or a mix of all three. (I recommend [Bret Deveraux's blog series](#) for a sense of what pre-modern farmers are likely to value and why.) Above all, though, it's a classic example of an "[illegible](#)" society: one that makes sense internally but is hard for centralized government to understand or extract from. Figes mentions a couple of occasions [p. 86, p. 136] when urban intellectual types have the bright idea of going "back to the peasants" for their political program. They usually wind up back in the cities within a year or two, whimpering in horror.

The awkward truth is that the peasants, a large majority of Russians, don't have a political program. Not in the sense of being internally divided, but in the sense of actively not giving a crap [Figes p. 581]. Their primary political stance is "leave us alone", which is understandable considering their situation and history. Their secondary political stance is "kick out the nobles and take their land", also understandable. Other than that, the only message that seems to mobilize them is nationalism leaning on religious and xenophobic appeals (which predictably led to anti-Semitic pogroms). In short, Russia's peasants aren't a major source of political leaders or ideas, but only of political power-- which in their context mostly means violence.

Everything comes to a head in February 1917. The war is going badly and that, combined with socialist resistance, has increasingly sapped Tsarist authority. The government can't effectively extract food from the peasants, transport and industry are breaking down [Figes p. 298], and the factory workers and soldiers increasingly answer to the Soviets (socialist workers' councils). Bread riots in Petrograd escalate into a general mutiny of the garrison, overthrowing the national government. The military persuades the much-despised Nicholas II to abdicate to restore order; Russians joyfully embrace freedom and the Duma negotiates the Soviets' support for a liberal-socialist coalition government, with democratic elections to follow in a few months.

Then Vladimir Lenin and John Reed show up.

### III

*Ten Days That Shook the World* opens with Reed describing the political situation in Petrograd in September 1917. It's not pretty. The provisional government has been ineffectual under the liberal noble Lvov and then the right-wing socialist Kerensky. Elections haven't happened yet, the war is going badly, and reforms have stalled. To make matters worse, through a series of boneheaded political maneuvers Kerensky has raised suspicions that he's plotting a military coup. The Soviets that *de facto* control Petrograd still have no direct role in the government and are dominated by "moderate" socialist factions-- Reed never fails to include the scare-quotes-- that lack the political will to push the revolution further.

(Figes [p. 331] gives the hilarious explanation for the Soviets' inaction: many moderate socialists refused to take political power or enact social revolution at this point because it contradicted their Marxist doctrine! Marx's theories focused on industrialized, capitalist societies like England and Germany. Russian society was still mostly pre-industrial and nobody thought workers could build capital on their own. So the plan was to put bourgeois capitalists in power, wait for them to industrialize, then have a second revolution and take their stuff. The same situation also led to ironies like Marxists pushing industrialization against the old-fashioned, agrarian-minded Tsarists [p. 119].)

Enter the Bolsheviks. They're a radical Marxist party that's come to prominence since February by demanding immediate "Peace, Bread, Land" on behalf of the soldiers, workers, and peasants respectively. The provisional government is unwilling to grant these demands in the current situation and the moderate socialists who control the Soviets are unwilling to force the issue, so the Bolshevik call for the Soviets to take power picks up a lot of support among the soldiers and workers. (The peasants are, as usual, a somewhat different story. The Bolsheviks don't really represent them-- the radical and moderate Marxists among them have their own parties-- and they're already getting the land they want by taking it from the nobles in small local uprisings. In Reed's telling, the Bolsheviks come off as constantly anxious about the support of the peasant representatives.)

Reed, repeating Bolshevik claims, presents the political and economic situation by October as an urgent crisis. Figes [p. 470] is skeptical. A national Soviet Congress, due to meet shortly, was likely-- given the provisional government's failures-- to take power peacefully and on its own terms. Such a broad socialist coalition would probably exclude Lenin, the most prominent Bolshevik leader, who was uncompromising and openly authoritarian in both his intra-party politics and his vision for Russian government. Lenin convinced the Bolshevik leaders to plot a violent takeover instead. Reed accepts his justification for urgency-- fear of a violent crackdown by Kerensky-- but Figes argues that the real reason was Lenin's desire to seize control for the Bolsheviks alone, and for himself.

To their credit, the other socialist parties understand exactly what's going on and have a brilliant series of talking points to oppose it. All of them, even the Bolsheviks, live in fear of "counter-revolution", i.e. supporters of the old government taking up arms and turning the revolution into a large-scale civil war. So they accuse the Bolsheviks of being "provocators", traitors out to force a violent confrontation that will cause counter-revolution and give reactionaries an excuse to destroy what the February Revolution gained. As it turns out, the prediction of counter-revolution isn't far wrong. But it isn't enough to stop the uprising.

The coup itself turns out to be ridiculously easy. Most of the Petrograd garrison sides with the Bolsheviks, and Kerensky flees the city, leaving only a few thousand pro-government holdouts in the Winter Palace, who melt away after ten thousand or so Bolshevik supporters march up. The harder part is staying in control. There's fighting in the streets. Most of the old government bureaucracy has quit in disgust. Large contingents from both the provisional government and the Soviet Congress-- which has finally assembled-- are protesting the Bolsheviks' legitimacy and challenging their authority. Kerensky has shown up again and is moving toward the city with fresh troops, and elsewhere in Russia there's even more resistance to the new government.

But the Bolsheviks weather all of this, more or less. Kerensky's troops are no more reliable than the Moscow garrison, and his push collapses. The moderate socialists are all bark and no bite. A scorched-earth takeover of government functions, and a series of repressive measures by the Bolshevik militias, restores a semblance of order. With their control of the government more or less an established fact, Lenin and company are able to strong-arm the Soviet Congress and a later Peasants' Congress into legitimizing a Bolshevik-majority government. (Reed presents these as mainly achievements of rhetoric, political maneuvering, and decisiveness, but Figses says there was undemocratic foul play involved as well.) The Bolsheviks then proceed to set up a Soviet government and pass a bunch of decrees enacting their "Peace, Bread, Land" program, finally completing the Russian Revolution. The end.

Confused yet? Just be glad you didn't have to read *Ten Days*. Reed has a bad habit of mentioning people, parties, and other features of Russia's political landscape without introducing or explaining them adequately. He's also prone to quoting official proclamations at length just because he can, and it's easy to get lost in his blow-by-blow account of events. In his telling the whole story becomes a soup of different factions, councils, and committees with politicians scurrying among them at random. Actually that might be what it looked like to Reed; he never makes this clear, but his access to the socialist leaders was pretty limited. In that sense *Ten Days* is more of a primary source document, like a diary or newspaper, than a work of history.

In other senses, though, Reed is trying to be a secondary source, an interpreter of events-- and this is where his writing calls for some untangling. Here's a typical passage from the first chapter: "For months in Petrograd, and all over Russia, every street corner was a public tribune. In railway trains, street-cars, always the spurting up of impromptu debate, everywhere...." The ellipses are Reed's, not mine. After the scare-quotes they're his favorite weapon of punctuation. They make his descriptions feel impressionist, like he's inviting you to fill in further instances and

details-- which he is. In fact a lot of these sweeping descriptions of his seem to be based on a few anecdotes. He ends up assuming, or generalizing, trends from not very much data at all.

Sometimes Reed shifts into the first person, and when he does it's a treat, because he's a good reporter, with a flair for vivid descriptions and a talent for getting into places he shouldn't be. He's at the Winter Palace a few hours before it falls, at the prison holding officials from the old government the next day, and at the front lines outside Petrograd when Kerensky is closing in. My favorite passage is the one where Reed tags along with a group leaving Petrograd and they're stopped by a couple of soldiers who support the Bolsheviks. The soldiers don't recognize Reed's pass from the socialist leadership, though, and they can't read, so they decide to shoot him on the spot. He only escapes by getting a local woman to read his pass to them. If the entire book were just these first-person sections it would be a lot more fun and interesting to read.

Most of the time, though, Reed is writing in the third person-- and according to Taylor's introduction, this often means he's repeating the Bolshevik version of events and not being up-front about it. This isn't just mild bias either; he gets important facts wrong. The worst howler is when he describes a series of drunken riots caused by soldiers breaking into the Winter Palace's wine cellars, and adds with a straight face: "In all this was evident the hand of the counter-revolutionists, who distributed among the regiments plans showing the location of the stores of liquor." Figs [p. 495] cites hard documentary evidence for the opposite, obvious conclusion: that the actual instigators were members of the badly disciplined Bolshevik force that had just captured the place.

## IV

Reed states the main thesis of *Ten Days That Shook the World* at the end of his second-to-last chapter. "Not by compromise with the propertied classes, or with the other political leaders; not by conciliating the old Government mechanism, did the Bolsheviks conquer the power. Nor by the organized violence of a small clique. If the masses all over Russia had not been ready for insurrection it must have failed. The only reason for Bolshevik success lay in their accomplishing the vast and simple desires of the most profound strata of the people." In other words, he sees his story as supporting two main claims. First: the Bolsheviks succeeded due to radicalism, not compromise. Second: they succeeded through popular support, not a military coup.

The scary thing is, there's a sense in which both claims are technically correct. Lenin's uncompromising radicalism was a critical part of the October Revolution, and it did hinge on the Bolshevik program satisfying... well, I would say the second-most-profound stratum of the people, since they didn't really represent the peasants. The problem is that in framing the issues this way, Reed is pushing an upside-down causal model of how the revolution happened.

When I first read about Bolshevik political tactics, I was so confused that it felt like I'd stepped into some kind of mirror universe. Against all political common sense they insisted on extreme positions, picked fights with close allies, and intentionally made crises worse. Even crazier, it seems to work for them. They're called the Bolsheviks ("Majority-ers") because they secured a majority in a party schism in 1903-- but only after allies of the opposing Menshevik majority staged a walk-out [Figes p. 384]. Reed, too, describes walk-outs from moderates paving the way for the Bolsheviks to secure power in 1917. Between those dates they were a minority, often a tiny one, in every political arena they entered; as late as 1915 they had fewer than 500 party members in Petrograd [Figes p. 297]. Yet they won!

How? I'm not sure of the exact explanation, but it all seems to come back to the weakness of Russian democratic values and institutions. First of all, that explains the frequency of walk-outs: walking out of a congress is basically a bet that your supporters have more loyalty to you than to the institution of the congress. Second, it speaks to the thinness of the middle classes-- the managers and intellectuals who might have stood up for democratic procedure; Figes [p. 520] describes this as simply "the tragedy of 1917". Finally, and related to the first two, one could describe the problem as a shortage of political legitimacy. The monarchy had soaked up all of it and staked it on a losing bet on pre-modernism. Once the Tsar abdicated, there wasn't enough of it left to sustain good-faith political actors, and not enough time to make more.

Reed focuses on a version of that third explanation. In his telling there are basically two kinds of political actors other than the Bolsheviks. First, there are socialists like the Mensheviks, whom he portrays, pretty accurately as far as I can tell, as idealists who dither ineffectively as events overtake them. (He describes a brilliant comic moment, on the night of the revolution, where they try to halt the Bolshevik forces with a nonviolent protest but end up losing their nerve and walking away.) Second, there's everyone to the right of the Mensheviks-- from monarchists to liberals to socialists like Kerensky. Reed casts them all as forces of reaction, either openly or secretly committed to undermining and rolling back the Revolution.

To some extent, even that description is true. The Kadet Party, a faction of upper-class liberals that was on the far left in 1905 [Figes p. 207], was on the far right by October 1917 and suspected of being secretly monarchist-- and in the Civil War it sided with the counter-revolutionaries. But its actual position hadn't changed much. What happened was an extreme political shift, and with it, polarization [Figes p. 457]. In part that was driven by popular dissatisfaction with the Tsar and, later, the provisional government; but it was also actively instigated by Lenin and the Bolsheviks with their ultimatums, their violent uprisings, and their "Peace, Bread, Land" slogans. Intentionally or not, they're following a strategy optimized for creating conflict-- because in a situation of conflict, compromise becomes unnecessary and even suspect, and radicals get empowered. Stop me if you've heard [that one before](#).

But aren't the Bolsheviks just expressing a fundamental need and desire of the common people? Figes isn't so sure. He repeatedly argues that the peasants in particular didn't really have political consciousness on a national scale. Instead they had a "parochial" political

philosophy, independent and traditionalist, which didn't necessarily align with the Bolsheviks'. What tipped them toward the Bolshevik side was exactly the polarization. During the Civil War peasants faced coercion from both the Bolshevik "Reds" and the opposing "Whites", but considered the Reds "the less bad of two bad options" [Figes p. 668] because on the whole they liked the revolution. After all, it had given them land.

Which brings us to Reed's claim of popular support. Retroactively, yes, the Bolsheviks had the support of the people-- but only because their brutal, violent, rigid, authoritarian movement succeeded in making itself the only viable alternative to the Tsarist social and economic system. If that's enough to call their rule democratic, it's hard to say when between 1917 and 1991 it stopped being democratic. The premise of elected government is that the party with an agenda the public prefers can win, right? And for all their ineffectiveness, it's fairly clear that the liberals and moderate socialists represented the preferences of the Russian people better than the Bolsheviks did.

Electoral-politics nerds already know the punchline here: [Arrow's impossibility theorem](#). Roughly, no matter how you set up an election, there's no way for the outcome to fairly reflect everyone's preferences. In that sense, there's no such thing as popular will; it's all a matter of framing. The October Revolution, though, takes that about ten steps further. It shows that in practice, even a really extreme and horrible framing can win out. All it takes is for the Tsar to insist stubbornly on one extreme, and then for the Bolsheviks to insist stubbornly on the other. If support for real democracy isn't strong enough to resist, polarization and political maneuvering will do the rest. In Soviet Russia, outcomes of elections decide YOU!

## V

Alexander Solzhenitsyn, under the Soviet regime, wrote that the line between good and evil runs through every human heart. I'm tempted to conclude something similar about the line between left and right. Political will isn't an inherent thing; it's situational and inconsistent and hard to model. Partisan politics, though, can't help flattening it into a cartoon. Maybe we should think of every human heart as a Soviet Congress of its own: moderates trying their best to hold it together as extremists on both sides obstruct, threaten walkouts, and wait for their chance to grab power.

At any rate that's how I'd like to think of John Reed. He was a skilled journalist and an energetic writer and probably a decent person-- not a would-be totalitarian at any rate. He could have written a truly gripping and revealing book about the Russian Revolution if he hadn't been so busy trying to stay on the right side of history. But he, like the Russian people, got taken in and then strong-armed by the Bolsheviks' political framing; and like the Russian people, he didn't have the habits of fairness and deliberation he needed to resist it. The book he ended up writing was deceitful, even though Reed himself wasn't.



Reed thought the story of the October Revolution proved that Communism was a fundamental, inevitable, necessary social change. After reading Figes, I think it proves the opposite. Studying history-- especially Russian history, I guess-- can easily end in fatalism: the last part of *War and Peace* is basically Tolstoy insisting that Napoleon's 1812 invasion of Russia had to happen as it did regardless of what anyone decided. But the Bolshevik coup is a relatively clear case of contentious, history-making, yet very personal decisions that could easily have gone either way. If Kerensky (better yet, his predecessor Lvov) had shown a little more initiative, or Lenin a little less, the history of the 20th century might have looked very different.

There's a quotation from Margaret Mead that I know by heart thanks to seeing it so many times in high school: "Never doubt that a small group of thoughtful, committed citizens can change the world. Indeed, it is the only thing that ever has." I did doubt that when I first heard it. I still doubt it in general. But reading Reed and Figes has given me new appreciation for just how much a small, committed group of citizens can do to change the course of a major world power for generations, defying all apparent principles of democracy in the process. Indeed, one might argue it's the only thing that ever did.

One final note: I researched and wrote this review on and off between May 2020 and February 2021. Since you're probably wondering, I'll admit that, yes, studying the Russian Revolution during this particular period was quite the eye-opening experience. I'm reluctant to make direct comparisons, not because they aren't valid but because so many different ones are; by committing to one analogy I might be dismissing another, even better one. Also, I don't want to come off as alarmist: there are plenty of ways 2021 America is in much better shape than 1917 Russia was. But if I've learned anything from *Ten Days That Shook the World* and *A People's Tragedy*, it's not to take those benefits for granted.

# The Age of Entitlement: America Since the Sixties

## by Christopher Caldwell

I wonder if Christopher Caldwell, author of **Age of Entitlement: America Since the Sixties**, had the same trouble titling his book as I did with this review. The most precise, encompassing title I can think of is, "How the Civil Rights Act Ruined America," but I'm worried that even putting that phrase at the top of a blog post will forever label me a Richard Spencer-style racist, so like hypothetical Caldwell, I played it safe.

Caldwell's problem with the Civil Rights Act isn't racial, but legal. By his interpretation, and supposedly the interpretations of lots of prominent legal scholars and judges prior to the modern era of PC dominance, the Civil Rights Act tore a fatal wound in the Constitution which has only festered over the last half-century. The core principle of the American Founding Fathers – not structural racism, but legal equality – was flipped on its head by a novel collection of legal precedents that formed a new "de facto Constitution." The result over the last sixty years has been a ballooning federal government, the erosion of individual rights (as opposed to "civil rights"), a "racialization" of the population, and a deepening cultural malaise for all Americans who aren't part of the "ascendant" coalition of wealthy liberal elites and their minority allies (which is not all minorities), at least until Trump.

Ok, that previous paragraph will definitely read as barely concealed alt-right racist bloviating to a lot of people. But I don't get the sense that Caldwell (not to be confused with actual racist, Christopher Cantwell) is a white supremacist or anything like that. Caldwell doesn't actually blame minorities or really anyone for the errors of the sixties. He portrays the Civil Rights Act, the feminist movement, the gay rights movement, etc. as being pushed by a bunch of well-intentioned but short-sighted activists who tried to make America better, and arguably did so for a significant portion of the country, but who also accidentally triggering a whole bunch of unintended consequences that made America worse as a whole. This doesn't mean that Caldwell wants to bring back segregation, send women back to the kitchen, and send gays into the closet, etc. Rather, he wishes these liberations were achieved in a more judicious manner with fewer second-order effects. The whole book reminds me of a really good Reddit post you might find on r/bestof or r/themotte. It's well-written, elegant, proposes lots of fascinating ideas and trends, but it also reads a lot like storytelling. On a big Reddit post it's generally accepted to throw out a bunch of conjecture for the sake of discussion, and it's usually fine to not cite much research or data since you're just having fun with ideas and want to see how they bounce off other Redditors. But **Age of Entitlement** seems to take this approach for the entire book. Caldwell makes massive sweeping assessments of the general political and cultural mood of the country, and usually doesn't offer more than quotes from key figures and occasional survey data to back them up. He also makes lots of huge causal leaps between disparate political and cultural trends, often with little attempt to justify the connections beyond just-so reasoning.

**Age of Entitlement** offers a radically different interpretation of the last sixty years of American history than the mainstream view I learned in school. By my evaluation, Caldwell's story is plausible though I am by no means committed to it. At the very least, I think the book offers a better explanation for a few key trends than the conventional historical telling, including:

1. The breakdown of Constitutional authority from the bedrock of American law to mere argument fodder
2. The highly ambiguous impact of the Civil Rights Act, and the massive taboo

around even suggesting that it had drawbacks or unintended negative consequences 3. The current (bad) state of American race relations 4. The ever-increasing size of the government, especially government spending 5. The accepted reality of the court system operating as an unofficial legislative body 6. The polarization of politics and society

This review is my best attempt to state Caldwell's position. I'd estimate that about 70% of the explanation comes straight from Caldwell, while the other 30% consists of my additions to glue the explanation together.

## I.

### Unifying Force

According to Caldwell (I'm going to be using this phrase a lot), all countries need a unifying force. For most, that force is some combination of ethnicity, language, religion, and shared history. But due to its fractious immigrant population, the United States can't use most of these tools. Instead, the US is one of the very few countries in the world that is primarily unified by *ideology*. In lieu of a common bloodline or ancient history, we found our bonds in the American Revolution and the Enlightenment ideas upon which it was based. The ultimate personification of this ideological unifier was, and still is, the Constitution, the fundamental legal and existential bedrock of the United States.

What are these Enlightenment ideas? They're notoriously hard to summarize, but I'd call them *a tendency towards negative individual rights on a classical liberal basis*. Meaning, all individuals have the right to live their lives as they see fit without interference from anyone else, including the government. That is, except when the government needs to do something super important for the continuity of the country. So the government can't tell me what not to say, can't tell me who I can't associate with, can't imprison me unless I have committed a crime, can't force me to quarter soldiers, etc. The government can tax me to maintain the army and roads to keep the country protected and operational, but not to finance a religion, or support an industry, or for economic redistribution, etc.

But the Constitution is not perfect at enumerating these ideas, nor were the Founding Fathers by any means unified on them. The Constitution was born out of compromises between the federalist and antifederalist wings of the Founding Fathers which debated the optimal degree of centralization in the American government. Over almost two hundred prosperous but tumultuous years, America struggled to figure out how to stabilize the government obligations and citizen duties described in the Constitution. The nation fought over taxation, national banking, imperialism, and more direly, slavery. But, according to Caldwell, there was always an underlying deference to the fundamental Enlightenment principles of the Founding Fathers. It was always the federalists vs. the antifederalists, albeit with different names over time.

In 1887, a severe drought hit Texas and wiped out the prosperous cattle industry. Congress convened and passed the Texas Seed Bill to provide \$10,000 to struggling farmers. President Grover Cleveland vetoed the bailout, and wrote to Congress:

"I can find no warrant for such an appropriation in the Constitution, and I do not believe that the power and duty of the general government ought to be extended to the relief of individual suffering which is in no manner properly related to the public service or benefit. A prevalent tendency to disregard the limited mission of this

power and duty should, I think, be steadfastly resisted, to the end that the lesson should be constantly enforced that, though the people support the government, the government should not support the people.”

More than 30 years later, President Herbert Hoover held the line against government welfare during the greatest economic panic in American history. Both presidents resisted breaking political traditions for the sake of momentary expedience and relief, even in the face of enormous populist pressure.

Such principled stands for Enlightenment values and Constitutional governance died with Hoover. The Progressive Movement and then FDR’s New Deal argued for the government to take on new roles which *fundamentally* altered the relationship between the government and the individual. In the 1930s, FDR introduced taxpayer-funded social security and unemployment benefits which *obligated* the government to provide money to its citizens. Now Americans were given material benefits from the government rather than just protections. This new trend pushed the US government towards *positive rights*, which Caldwell calls *entitlements*. Such measures were not entirely unprecedented in American history, nor did they constitute a significant component of government responsibilities, nor of the American economy. To Caldwell, they were just more awkward, contradictory chunks of American governance held up by the underlying Founding Fathers principles.

However, everything would change in the 1950s when the civil rights movement would lean on these contradictions until the foundations beneath them collapsed.

## The Worst Court Case in American History

In 1951, civil rights leaders threw their weight behind *Brown v. Board of Education*, a class lawsuit by representatives of twenty black children and their parents in Topeka, Kansas against a public school enforcing racial segregation.

Yes, according to Caldwell, *Brown v. Board of Ed* is the worst court case in American history. No, he doesn’t think that because he’s a massive racist who wants to segregate black children. He thinks that because, according to Caldwell, the *Brown* decision was not made in accordance with fair-and-balanced legal proceedings. It was made by politically motivated judges with the support of a public manipulated by a crafty PR machine intending to circumvent proper legislative procedures to push a radical agenda. *Brown* was the first major successful salvo from the civil rights movement, and it created the breach through which the whole movement would charge through and destroy the Constitution.

## It’s All In the Game

In elementary school I was taught that Rosa Parks was a kindly old lady who wanted to sit at the front of the bus after a long, hard day at work. She was asked to move to the back of the bus to free up the best seats for white people, but she was so tired that she refused. Her mistreatment at the hands of the Montgomery bus system and ensuing arrest so incensed the public that it triggered a massive bus boycott and pushed America a step closer to ending segregation.

What I learned years later was that Rosa Parks had been a political activist for almost fifteen years prior to the famous bus incident. When she refused to get up from her bus seat in 1955, she was acting as an agent of the NAACP in the first stage of a planned city-wide boycott and

lawsuit. Parks was specifically chosen for the task because she was a little old lady who would garner sympathy from the public and thereby put pressure on the powers-that-be to oppose segregation.

Of course, this doesn't invalidate Parks's cause, nor her bravery (it may even buttress it). But it does make the famous arrest feel a bit... *inauthentic*. It shifts a spontaneous thrust for justice into a calculated, coordinated campaign with a public relations focus. The public was being sold something, even if it didn't know it.

According to Caldwell, this was the civil rights movement's modus operandi. Sympathetic figures didn't arise organically, rather, they were *orchestrated*, or at least carefully selected.

Orchestrating sympathy for the sake of PR in a broader culture war isn't necessarily a bad thing in and of itself. After all, there surely were countless black people regularly demeaned by racist segregation laws and policies across the South, just like Rosa Parks. It didn't hurt anyone to create a story that minimized complications and maximized sympathy for the just moral goal of ending segregation.

What troubles Caldwell about this strategy is its use in the justice system. According to Caldwell, prior to the civil rights movement's heyday in the 1950s, this orchestration strategy was considered *extremely bad form* by judges and juries. It was seen as a purposeful manipulation of the legal system. Judges are supposed to be objective and focus on the letter of the law, but everyone knows they can be swayed by public opinion and personal characteristics of plaintiffs and defendants. Orchestrating a case wasn't illegal, but it cynically went against the spirit of objective justice for which the legal system should be striving. Oddly enough, Caldwell either doesn't bring up or doesn't know that *Plessy v. Ferguson* used the orchestration strategy. *Plessy* was the first judicial attempt to outlaw segregation, but it tragically backfired and created the "separate but equal" clause. But *Plessy* happened all the way back in 1896, when civil rights and its associated movement arguably weren't a thing. Maybe the Supreme Court was too conservative during the following fifty years to use the orchestration strategy again? I don't know.

Either way, *Brown* was the first major modern civil rights case to use the orchestration strategy and set the precedent that judges were ok with manufacturing cases, even in the highest court in the land. *Brown*'s lawyers had combed the country for the perfect case against segregation, and they settled on literal poster children with minimized blemishes and maximized poise to elicit public sympathy... and it worked. *Brown* overturned an almost 60 year old ruling which had served as a lynchpin for pro-segregation law. In turn, nearly all of the major civil rights court cases – *Griggs v. Duke Power*, the Parks-based *Browder v. Gayle*, etc. – were based on highly selected or outright orchestrated representatives to sway judges and juries

## The Soft Underbelly of the US Government

So what if clever civil rights activists used orchestrated sympathetic cases to boost their odds of winning lawsuits? Ending segregation was a worthy goal. Even if the means were a little phony, all is fair in love and war and stopping literal systematic racism, right?

No, not according to Caldwell. Because the prosecution of *Brown* was based on a broader civil rights strategy to achieve policy goals in a historically unprecedented and unconstitutional manner, thereby eroding the very basis of American law.

Prior to *Brown*, the civil rights movement had been trying for decades to fight segregation legislatively in Congress, but never found success. The support for segregation was far too entrenched in the South, and the rest of the country was too apathetic to throw serious political capital behind it.

In the 1950s, civil rights leaders shifted to another strategy – targeting the courts. Getting Congress to pass new laws required tremendous lobbying, political gamesmanship, and swaying the entire American electorate towards a view. But winning in the courts only required one good case. One good case could break a law painstakingly passed by Congress, or, better yet, create a legal precedent that steered countless more laws. Hence, the judicial system was the backdoor into American governance.

Of course, there was one massive obstacle to this strategy – the Constitution. The whole original purpose of the judicial branch of government was just to read laws and decide whether they abided by America's founding legal document. The best cases and lawyers in the world wouldn't change a law via the courts if they supported unconstitutional objectives.

That is until the rise of the legal philosophy of judicial activism. Under the old standard of judicial restraint, the Constitution was treated as a fixed set of ideas against which novel laws could easily bounce off and die. But under the theory of judicial activism, the Constitution became a living document that needed to be adapted to the times. This allowed for more and more liberal interpretations of the Constitution and potential for major shifts in the accepted constitutionality of laws.

Fortunately for the civil rights movement, the Warren Court (1953-1969) was the most liberal, judicially active Supreme Court in American history and thus created a perfect setting for the movement to spawn cases and take shots at long-standing legislation and legal precedents. Enter *Brown v. Board of Ed.* Prior to the case, segregation had been declared legal by the "separate but equal" doctrine in the 1896 *Plessy v. Ferguson* case. The *14th Amendment* had declared legal equality for all classes of citizens in the US, but *Plessy* argued that segregating races didn't necessarily make them legally unequal, only... you know, separated. The Supreme Court ruled that this interpretation of the Constitution was accurate, regardless of one's feelings about racial equality.

*Brown's* 1954 decision, decided by the Warren Court, overturned *Plessy* and rendered all government-instituted segregation unconstitutional. Which sounds great! Everyone today agrees that segregation is wrong, and even diehard libertarians don't think the government has the right to discriminate based on race. So why would Caldwell have a problem with *Brown*? Because according to Caldwell, the *Plessy* decision was correct and the *Brown* decision was incorrect from a strictly Constitutional perspective. "Separate but equal" is perfectly fine from a literalist (ie. accurate) reading of the Constitution. That doesn't mean segregation is morally good, it just means it's constitutional.

According to Caldwell, *Brown's* side in the case was trounced by any reasonable legal standard. Yet *Brown* won because... well, the Warren Court wanted them to win. The extremely liberal, anti-segregationist Warren Court was severely biased against *Brown's* opponents, so they wrote up a bullshit legal justification to give *Brown* the win and judicially overturn segregation legislation.

Going outside Caldwell for a bit, it's difficult to understate how shocking *Brown* was to much of the US. I'd say the decision was akin to the Supreme Court legalizing gay marriage in 1995.

There was a general sense of societal movement away from segregation, but likewise there was a sizeable reactionary minority of the population which had tightened its grip on the institution over recent years. When government segregation was outlawed overnight, the South's response was maybe a degree below armed revolt, with states indefinitely shutting down public schools and governors ordering the national guard to enforce segregation.

I most certainly don't have the legal knowledge to evaluate Caldwell's claim, but some Googling and Wikipediaing indicates that it's plausible. At the time of *Brown* especially, there seems to have been a strong backlash against the decision from the legal field. From the Wikipedia entries on *Brown* and the Warren Court:

William Rehnquist wrote a memo titled "A Random Thought on the Segregation Cases" when he was a law clerk for Justice Robert H. Jackson in 1952, during early deliberations that led to the *Brown v. Board of Education* decision. In his memo, Rehnquist argued: "I realize that it is an unpopular and unhumanitarian position, for which I have been excoriated by 'liberal' colleagues but I think *Plessy v. Ferguson* was right and should be reaffirmed."

Chief Justice Warren's reasoning was broadly criticized by contemporary legal academics with Judge Learned Hand decrying that the Supreme Court had "assumed the role of a third legislative chamber" and Herbert Wechsler finding *Brown* impossible to justify based on neutral principles.

Some Constitutional originalists, notably Raoul Berger in his influential 1977 book "Government by Judiciary," make the case that *Brown* cannot be defended by reference to the original understanding of the *14th Amendment*. They support this reading of the *14th Amendment* by noting that the Civil Rights Act of 1875 did not ban segregated schools and that the same Congress that passed the *14th Amendment* also voted to segregate schools in the District of Columbia.

[Richard] Fallon says that, "Some thrilled to the approach of the Warren Court. Many law professors were perplexed, often sympathetic to the Court's results but skeptical of the soundness of its constitutional reasoning. And some of course were horrified."

If the *Brown* decision was blatantly unconstitutional, then it was essentially a power grab by the civil rights movement. A power grab with a morally good objective, but a power grab nonetheless.

Remember, to Caldwell the Constitution had been the unifying force of the United States since its creation. By callously ignoring the Constitution, the civil rights movement was striking at the very bedrock of the US government, culture, and sense of civic unity. With *Brown*, the civil rights movement won a victory, but at what cost?

## The New Constitution

If *Brown* tore a hole in the constitution, the 1964 Civil Rights Act tore it in half. This one piece of legislation changed America more than literally any other bill in history.

Yes, I just said that the civil rights movement always failed in the legislature and thus shifted to a judicial strategy, but Caldwell argues that the Civil Rights Act was the exception that proved

the rule. The bill only passed because civil rights were being championed by America's beloved, young, handsome, charismatic president, John F. Kennedy, who then took a bullet to the head in front of his wife and garnered an unprecedented outpouring of sympathy from the public (I'm told my Irish Catholic grandmother literally built a shrine to JFK in her living room). Kennedy's successor, President Lyndon Johnson was a notoriously smooth political operator who capitalized on the national mood to push the Civil Rights Act through Congress in Kennedy's name.

*Brown* only ended segregation in government institutions. The Civil Rights Act was a massive pile of new laws and regulations which sought to end segregation in private institutions. As any good libertarian will tell you, there's a world of difference between those two goals. The public schools targeted by the *Brown* case were using taxpayer funds to force students to abide by racist rules at an institution they were forced to attend. But the racist segregation rules maintained by thousands of restaurants, cafes, stores, social clubs, etc., were instituted on private property by American citizens. And in America, private property is sacred; the government isn't supposed to be able to force someone to surrender, alter, or use their private property in a particular way, except for the sake of the literal continuity of the nation.

Even if the *Brown* decision was Constitutionally unfounded, at least the ruling only negated a precedent derived from a post-Civil War Amendment. But the Civil Rights Act was a direct assault on the *1st Amendment*, or at least some of the earliest and most well-founded precedents that the American judicial system derived from it. The *1st Amendment* textually guarantees Americans the right to not have their speech censored by the government, not to be prohibited from practicing a religion, and not to be prohibited from assembling. But a long-understood implication of the amendment was the *right to free association*. That is, Americans can speak, meet, and interact with whomever they want, or avoid interacting with whomever they don't want to.

There was no way around it: *the Civil Rights Act violated the right of free association*. Its laws gave the government the power to force individuals to associate with people they didn't want to. A store owner couldn't choose who he sold to, a country club couldn't choose its membership, a vendor couldn't choose its supplier, etc. Again, the motives were good (to end systematic racism), but to Caldwell, the US government had never engaged in such a direct violation of individual rights outside of wartime.

One point Caldwell repeatedly hammers home is that *nobody* understood what they were getting with *civil rights*. Nobody even really knew what the term meant. Ending segregation was a noble goal, but seemingly neither the voters, politicians, nor even the civil rights advocates understood what the government would need to do to achieve that end.

Consider the 1971 case of *Griggs v. Duke Power Co.*, which I'm pretty sure Caldwell thinks is the second-worst American court case of all time. For decades, Duke's black employees had only worked in the company's low-paid labor department. It's not clear, to my knowledge, if this was due to actual racial bias or simply a coincidence. In 1955, Duke added a requirement that employees must have a high school diploma to work outside the labor department.

In 1965, the Civil Rights Act passed, and one of its many laws prohibited racial discrimination in hiring. The bill also had a provision that if a company gated jobs with tests or requirements (such as needing to have a high school diploma), the government could challenge the company to prove the requirements are "reasonably related" to the job.



Not wanting to get sued into oblivion, Duke hedged its employment structure by adding two tests to the application for non-labor department jobs: a mechanical aptitude test and the Wonderlick test (a form of IQ test).

Nevertheless, Duke got sued for hiring discrimination. From Wikipedia:

Blacks were almost ten times less likely than whites to meet these new employment and transfer requirements. According to the 1960 Census, while 34% of white males in North Carolina had high-school diplomas, only 18% of blacks did. The disparities of aptitude tests were far greater; with the cutoffs set at the median for high-school graduates, 58% of whites passed, compared to 6% of blacks.

Again, it's not clear if the Duke executives really were racists trying to keep blacks confined to menial labor, or if they were just trying to put employees where they were best suited and tried to avoid a lawsuit with the new tests.

Either way, Duke lost the lawsuit. Once again, Caldwell thinks this was a clear case of motivated judges ignoring textual law for a preferred outcome. Duke argued that having a high school diploma and sufficiently high scores on a mechanical aptitude test and IQ test were "reasonably related" to working as an executive, accountant, or engineer at a power plant. The Supreme Court's judgement compared this argument to the Fox and Stork Aesop's Fable, where a fox and stork attempt to drink from the same vase, but its narrow opening only allows the stork through.

The Griggs decision represented a massive shifting of civil rights goal posts. The Civil Rights Act already placed the burden on employers and business owners to prove they weren't racist. But now even providing objectively equal opportunities to all races wasn't enough to prove non-racism. An employer also had to guarantee that outcomes of its sales or employment were racially equal, regardless of how that might adversely impact the very function of the entity (ie. putting unqualified individuals into engineering positions) or impose costs on the entity (ie. forcing Duke to expend resources to search for qualified black engineers).

To Caldwell, the Civil Rights Act itself was just the beginning. A string of major Supreme Court cases over the following decade (most of which used the orchestration strategy) would expand and cement the bill's power. Griggs was just one of many cases that imposed radical new rules on the American legal system and trampled the old conception of individual rights. Eventually, this network of laws, regulations, and judicial precedence amassed to become a new "de facto Constitution," which replaced the old "de jure Constitution."

The old de jure Constitution was based on the Enlightenment-based individual negative rights conceived by the Founding Fathers. The new de facto Constitution was based on progressive positive civil rights and their resulting entitlements conceived by the civil rights movement.

## Reconstruction 2.0 - The Largest Undertaking in the History of the American Government

According to Caldwell, desegregation and the civil rights project was the largest undertaking in the history of the American government.

He doesn't say that flippantly or hyperbolically, he means it literally. He believes desegregation,

or *Reconstruction 2.0*, was a bigger marshaling of resources and execution of policy than the first Reconstruction, the New Deal, or fighting the Civil War, World War I, or World War II.

Granted, Caldwell conceptualizes “desegregation” a bit differently than most people. He argues that it started with *Brown*, exploded with the Civil Rights Act, entrenched with ensuing court cases, and then continued with affirmative action, the Great Society, the War on Drugs, the increasing government regulatory burden on the economy, rapidly expanding national debt, and the government-encouraged ballooning of financial credit.

To take a step back – civil rights activists told a whiggish historical story of America starting as a noble ideal of equality muddled by systematic racism and sexism. But over time, America had shed its ignorance and slowly raised more of its population to full rights-bearing citizenship. The civil rights movement positioned itself as the next giant thrust in this story – the end of de jure racism and the bringing of America’s black population up to equal status with whites.

To achieve racial equality, the movement needed the support of non-Southern whites to use federal power to force civil rights on the South. But in Slate Star Codex terms, Southern blacks were a *far group* to non-Southern whites. These whites sympathized with blacks and their struggle against oppression, but this injustice was happening on the far side of the country where everything was weird and culturally different anyway. Northern whites didn’t understand in any deep, meaningful sense what segregation was or what would happen when it ended. Most assumed that ending segregation would just stop police officers from smashing black people over the head with batons in the street, and stop racist restaurant owners from enforcing their petty service discrimination.

But Caldwell these Northern whites were well-intentioned, but short-sighted.

An analogy - when ending slavery was debated in mid-19th century America, one of the main points of contention was what would happen to slaves once they were freed. Yes, they’d be free, and that was great, *but then what?* Through the worst form of systematic oppression, 4 million ex-slaves would start their free lives undereducated and impoverished. Could they survive on the labor market? Would they be paid a fair wage? Would they turn to criminality? Reconstruction was an attempt to solve these problems and start free blacks on a stable footing, but a largely unsuccessful one.

Caldwell’s argument follows a similar line of reasoning. The oppression of blacks under Jim Crow wasn’t as bad as slavery, but American blacks were still massively disproportionately undereducated, impoverished, and now in the penal system. Only outlawing segregation, as the white electorate thought they were doing with the Civil Rights Act, would be the equivalent of liberating slaves and then leaving them to their own devices.

In other words, the Civil Rights Act needed a Reconstruction 2.0 – a set of laws and aggressive policy initiatives to support liberated blacks.

However, there is a fundamental difference between the Reconstruction of the 1860s-80s, and the Reconstruction 2.0 of the 1960s-present. The first was based on *legal protection*, the second was based on *civil* protection.

By the 1960s, the gains blacks could make legally were at their end. With the *Brown* decision, which said that the government couldn’t racially discriminate, blacks had attained full de jure legal equality with whites. Private actors could still discriminate, but only within the confines of

their own property rights. But clearly the country was not enjoying racial equality in 1954. Even if *Brown* was fully enforced on state and local governments (it wasn't), there was still a non-governmental regime of segregation across the South making society de facto racially unequal, regardless of de jure laws.

In other words, civil rights advocates had gone as far as they could toward achieving racial equality through traditional legal rights. If they wanted to push for more equality, they had to go past the law and into *civil society*. That is, they had to figure out a way to change social structures rather than just government laws and their enforcement.

Thus the Civil Rights Act and the ensuing Reconstruction 2.0 project was a series of attempts by the government to use the power of the state *to take control of civil society*. It was no longer enough for the government to make laws based on protecting people's old-fashioned rights of property and safety; now the government was obligated to provide people with *entitlements*. Specifically, racial minorities were given entitlements (affirmative action, welfare, etc.) meant to boost their social status to close the racial gap.

One of the first major entitlements the Civil Rights Act created was the right to service and employment from *private parties*. The Act made discrimination illegal even within all other proper legal boundaries, like one's private property. And with the Griggs decision, even non-discriminatory actions were illegal *if their result resembled theoretically discriminatory actions*. The old right of an individual to use his property as he saw fit was trumped by the new entitlement of other individuals to derive benefits from someone else's property regardless of the wishes of the owner.

The Great Society and the War on Drugs were the next stages in the government's attempts to steer civil society after the Civil Rights Act. The Great Society was framed as a "war on poverty," but of all the groups in America, which was most impoverished? Blacks. The introduction of Medicare, Medicaid, Social Security expansion, and a huge increase in domestic wealth transfers were all intended to improve the material conditions of blacks to get them closer to racial equality. Plus [Medicare](#) "spurred the racial integration of thousands of waiting rooms, hospital floors, and physician practices by making payments to health care providers conditional on desegregation." The War on Drugs was a more even more misguided effort, but nonetheless was a well-intentioned attempt to rid black communities of a menace connected to poverty, crime, and undereducation.

The idea that these massive welfare programs were really a form of racial reparations might seem like a big stretch. According to Caldwell, pretty much *every major domestic policy* the federal government conducted after the Civil Rights Act *all the way up to the present day* has been an attempt to bring American blacks up to the level of American whites, or a consequence of it.

This is one of many areas where Caldwell gets very fuzzy and doesn't do a great job drawing the causal arrows. But by my understanding: Of course neither the Great Society nor War on Drugs were explicitly targeted at blacks in the way the Civil Rights Act was. And Caldwell doesn't suggest there was some sort of conspiracy to trick the American people into giving their tax dollars to blacks, or anything like that. Rather, he frames it as something like a difference between *de jure rights* and *national focus*.

Prior to the Civil Rights Act, only white Americans had full political rights, and therefore had full national focus, or the attention of the government. The American poverty rate was fairly low,

*especially if you only counted whites*. Black poverty rates had been declining at least since WW2, and quite quickly at that, but were still 4X higher than white rates by the Civil Rights Act.

Once blacks were given full political rights by *Brown and the Civil Rights Act*, there was suddenly a huge new portion of the American polity living in abject poverty. The rate of poverty amongst active participants in the political process spiked even though white poverty rates hadn't changed. This gave far greater impetus to the American electorate's default concern for poverty, and therefore inspired the passage of the welfare programs in the Great Society. A similar argument could be made about the (perceived) impact of drug use on black vs. white communities and the War on Drugs.

The Great Society, War on Drugs, and a host of other minor domestic initiatives would have been unthinkable before the 1960s. America was the last major Western nation to adopt government welfare programs, and only did so on a small scale in the depths of the Great Depression. But by the 1960s, the federal government was running a large wealth redistribution system despite a booming economy and the country's emergence as the preeminent global power.

Once again, Caldwell describes these policy shifts as the creation of entitlements. Americans no longer expected the federal government to merely fund a military and court system on 3% of GDP. Rather, Americans became *entitled* to unemployment insurance, social security, medical insurance, food stamps, drug-free neighborhoods, and eventually cheap credit.

Putting aside the national mood and contextual circumstances, the most fundamental reason why America never had a Great Society, War on Drugs, or Civil Rights Act before the 1960s was that they would be considered *unconstitutional*. I don't just mean that in the legal sense, but in the broader *spirit* sense. These policies went against the principles of the Founding Fathers and Enlightenment-based negative rights. Only a blatant reinterpretation of the Founding Fathers and a philosophical turn towards positive rights and entitlements could permit them. But that's exactly what the civil rights movement did – it ended the age of rights protection and began the age of entitlement.

Caldwell asserts that this fundamental restructuring of American rights was *not* the intention of the American electorate in the 1960s and onward. It may have been the intention of civil rights leaders and some political elite allies, but Caldwell is unclear on that. Though he doesn't use the term, *slippery slope* seems most apt. Americans were on board with giving blacks legal rights, but once blacks were full citizens, they had to be compensated for injustices, and materially provided for, and protected from harmful drugs, and so on. This resulted in countless massive legislative projects and court cases running through the system which irrevocably reshaping the United States. No one drove this whole process, it just happened once the ball got rolling.

## The False Prophet

According to the orthodox political understanding of American history, the hippies took over America in the 1960s and the US went from the austere WW2 mindset of the 1940s and 50s to the free-loving, soft governance of civil rights, welfare, and entitlements. But after this euphoria broke down into malaise in the 70s, Ronald Reagan rallied conservatives across the country, especially the long-suffering middle-class everymen of the "moral majority," and closed the door on all that hippy nonsense with his presidency in the 80s. Reagan made people proud to be American again as he stood up to the Soviets, cut taxes, brought capitalism back, and halted (though didn't reverse) the ever-radicalizing social goals of progressives.

Caldwell claims this view of Reagan is half right and half completely inverted. Reagan did indeed lead a coalition of conservatives against the progressive tide, and he got them some genuine wins in government, but in doing so, Reagan actually *entrenched* the civil rights movement and Reconstruction 2.0 while further undermining the Constitution and the original American vision.

According to Caldwell, it's possible that Reagan and his advisers really did try to roll back many of the reforms of the 1960s, at least at first. They didn't try to end the Civil Rights Act (the American people had accepted it fully by the 80s), but they may have tried to cut the ever-expanding budget deficit, cut the Great Society programs, and scale back economic regulations. There were even attempts to halt affirmative action and some of the more over-reaching aspects of anti-discrimination laws by appointing conservative judges.

However, Reagan soon discovered an intractable law of political science – it's vastly easier to create an entitlement than end one. After fewer than two decades, Medicare, Medicaid, and the other welfare enhancements were thoroughly entrenched in the economy. Taking money away from poor people was a political non-starter. Going after affirmative action, even through the courts, was equally fruitless in the early days of political correctness. But Reagan had to give his constituents something. He couldn't stop the entitlements of the progressives like he promised *but he could give the conservatives their own entitlements*.

Thus Reagan pushed the first significant tax cuts since the end of World War II. But the difference between these tax cuts and pretty much every other tax cut in American history up to that point was that they weren't matched by spending cuts. On the contrary, spending increased significantly (from 20.1% of GDP on average in the 1970s to over 21% in the 80s despite a rapidly expanding GDP) while revenue net-stagnated across Reagan's terms. The result (from Reagan's Wikipedia):

During Reagan's presidency, the federal debt held by the public nearly tripled in nominal terms, from \$738 billion to \$2.1 trillion. This led to the U.S. moving from the world's largest international creditor to the world's largest debtor nation. Reagan described the new debt as the "greatest disappointment" of his presidency.

Reagan was letting the American people have their cake and eat it too. They could have a safety net and massive military, but also low taxes and a roaring economy. Everybody wins! Except whoever has to pay for all this stuff in the future.

Thus to Caldwell, Reagan's reforms were an extension of the age of entitlement. His novel addition was the creation of entitlement to a high-spending government without having to pay for it with taxes. And like all entitlements, this policy was super easy to implement and virtually impossible to stop, hence almost every regime since Reagan's has increased its deficit spending, and there is no end in sight to the fiscal irresponsibility.

Reagan's worst crime of all was that he may have permanently cemented the age of entitlement into the American government and culture. Essentially, the Reagan reforms signed a peace treaty between progressives and non-progressives. The treaty said (in an abstract, unwritten manner) that progressives could keep all their civil rights reforms and welfare as long as they gave everyone else tax cuts and pro-American rhetoric. Reagan's regime was the last, best hope to halt the perpetual undermining of the Constitution in America, but instead he made peace with the enemy and reinforced their errors.

This peace treaty more-or-less held up until the disgruntled masses were so pissed off by collaborating elites and the illusion of choice that they thrust a certain populist demagogue into the most powerful office on earth.

## Pandora's Box

Everything above this point explains Caldwell's central thesis with a lot of my own framing and arguments acting as the mortar between points. Basically, the civil rights movement expanded political consideration to blacks, and this triggered a fundamental restructuring of the American foundation for political rights which promoted 50+ years of increasing entitlements for the American people. But I don't want to give the impression that Caldwell's analysis focuses only on American blacks. They were the focal point of the Civil Rights Act and Caldwell seems to believe they are still the focal point of progressivism today, but he argues that a similar trend occurred with other groups, especially women, gays, and immigrants.

I won't go into these trends too much, but the basic idea is that these groups used the ideological breach forged by the civil rights movement to first gain their own political status, and then push the government to try to steer civil society to their own ends.

By my reading, a lot of these causal chains are even fuzzier than Caldwell's connections between the civil rights movement and individual Reconstruction 2.0 policies. All three groups utilized the anti-discrimination framework created by the Civil Rights Act and its ensuing court cases for their own benefits. Women especially used a lot of the same rhetoric as blacks in the civil rights era to increase their power in civil society. Immigrants got their big push with liberalization in the 1980s as part of Reagan's compromise. And gays copied the best practices of the successful civil rights campaigns of blacks, women, and immigrants to craft their own uber-efficient civil rights effort in the 90s and 2000s.

These smaller civil rights movements all forced the government to give these groups their own array of entitlements. Women attained gains in abortion and divorce settlement. Immigrants got multiple amnesty rounds. Gays are the most recent and haven't gained as much as the others, though Caldwell makes an interesting (if somewhat conspiratorial) argument that gay marriage was primarily driven by wealthy gays trying to avoid inheritance taxes. All three groups got considerable anti-discrimination support from the courts, and are now considered protected classes.

## II.

### The Costs of Entitlement

Let's say you're not a hardcore classical liberal or libertarian, and you think sacrificing or reinterpreting the Constitution for the sake of giving blacks, women, immigrants, and gays equal rights is a worthy trade-off. And further, you think a welfare state and anti-discrimination laws are essential parts of modern governance, and not some weird, anomalous attempt by the government to forcibly reshape society that causes untold negative second-order effects. Given that the vast majority of Americans fall into this camp, why does Caldwell's analysis hold any water? Why should we care that legal norms conceived over 200 years ago have changed for the sake of equality?

According to Caldwell, you should care because this change has basically caused every major

domestic problem in America over the past 50 years. The legal system, culture, economy, and lifestyle of America has been permanently damaged to everyone's detriment. Before I describe how that's the case, I want to throw out the massive caveat that Caldwell isn't in favor of rewinding America back to the 1950s, and especially not to its segregationist regime. In fact, Caldwell has a whole chapter about how much he hates the 1950s not just for its racism, but for sort of militarizing America and priming it for failure. **Age of Entitlement** is a critique without proposed solutions. I'll try to list and summarize the cursed legacy of the 1960s. As with the above, this is a mixture of Caldwell's arguments and my own extrapolations of them, at about a 70-30 ratio.

## Racialization

One of the worst consequences of the age of entitlement is what Caldwell calls *racialization*. Essentially, Americans are highly aware of race in a political and civil sense, and have become ever more so since the 1960s to the point of obsession.

Caldwell claims that pre-1960s America was set on a trend of declining race consciousness. This seems hard to believe with the Southern segregationist regimes in full force, but it's plausible in the rest of the country where blacks had had far longer to integrate into normal society. Then in the 1950s, the civil rights movement purposefully raised race consciousness for blacks and whites alike to draw attention to the injustice blacks suffered. Whites who bought this message assumed that race consciousness would decline again after the Civil Rights Act as blacks were raised to equal status with whites, but instead the opposite occurred. The Act, its court cases, the Great Society, the Drug War, and all the other components of Reconstruction 2.0 rose in concert with race consciousness, though it's not clear which direction the causal arrow went.

This is a point that has been hashed out ad nauseum in anti-SJW circles, but it doesn't seem societally healthy to maintain high race consciousness. The entire point of the original civil rights movement and Martin Luther King Jr.'s arguments in particular were that whites and blacks were not fundamentally different and therefore should be treated the same. But cultural racialization ensures an obsessive focus on race which only serves to highlight evermore minute differences. As a result, what progress has been made in achieving civil racial equality is marginalized as goalposts are constantly shifted and new benchmarks made. To be clear, this is by no means entirely the fault of the civil rights movement and progressives. The reactionaries who opposed the civil rights movement most certainly fanned the flames of racialization and encouraged racial minorities to adopt id-pol stances. I think Caldwell actually downplays this a bit, and doesn't talk enough about the very strong, very real backlash to the civil rights movement in the 1950s and 60s and its enduring effects. I think we can even reframe part of Caldwell's thesis as *the civil rights movement kicked the hornets nest too hard with no tangible objective*. And that's how we end up with our confused, chaotic, exhausting present racial-sphere. There is no endgame. No solution. No compromise. There are just ever more racial battlefields to pit Americans against one another.

## White Malaise

A more recent unintended consequence of racialization is *white malaise*.

According to Caldwell, the white electorate has always been vaguely aware that the Civil Rights Act and its ensuing court cases gave non-white Americans a de jure and de facto advantage in the American legal system. For instance, we all know that a white-owned restaurant throwing

out a black patron would attract vastly more scrutiny and ire than if the races were reversed. And affirmative action is a form of literal institutionalized racism, both in its private use that's protected by the courts, and its use in public institutions.

Supposedly, this direct legal inequality was tolerated by whites because of a perceived inverted civil inequality. Whites had higher per capita incomes, lower poverty rates, more education, and less criminality, so letting blacks have special privileges in the legal system seemed like an ok compromise.

But over the last decade, that perceived paradigm has faltered. Yes, all those racial wealth metrics still hold when abstracted to the entire country, but a significant portion of the white population isn't feeling that advantage anymore. For the first time in pretty much ever, white American life expectancy is declining. The days of the crack epidemic in black inner-cities are behind us, and now we fear the white rural opioid epidemic. Deindustrialization has robbed huge swaths of the white rural countryside of its economic vitality. White labor force participation rates have plateaued.

Yes, the black population continues to struggle with poverty, undereducation, and criminality as much as ever in the aggregate, but there's a pervasive sense among many Americans that the *national focus* is still pointed at blacks at the expense of whites. Affirmative action is still permitted, and racial minorities get lots of special protections in the courts that whites don't. All the civil society intrusions created to promote blacks to racial equality remain while no reciprocated assistance is offered to mainstream whites despite their decline.

To massively compound this problem, political correctness and social justice have become deeply entrenched in mainstream American culture, particularly the media, entertainment, and academia. This phenomenon is a radicalized cultural outgrowth of the civil rights movement and Reconstruction 2.0. It has picked up where the court cases, Great Society, and Drug War have left off to enforce sweeping changes into civil society ostensibly for the sake of racial equality. But while state institutions are at worst ignoring the decline of white middle America, the SJW movement is exacerbating it by pushing American law and civil society into higher racial consciousness and more anti-white attitudes.

Caldwell *does* frame the white reaction to Reconstruction 2.0 and SJWism as a growing sense of white racial consciousness. But he portrays it as less of a simmering anger and more of a *malaise*. He sees a broad resentment growing in the white population, but not against blacks or other racial minorities, but rather against the coastal elites who harnessed the political and cultural trends started by the civil rights movement to entrench themselves as elites. The most extreme manifestations of PC we've seen in the universities and media is emblematic of this divide. Impoverished, undereducated, overincarcerated inner-city blacks aren't trying to get people fired for telling jokes, and they aren't whining about safe spaces. That's the domain of individuals of all races at Ivy League schools and powerful tech companies.

Supposedly that's why America got President Donald Trump. The white majority didn't elect a bona fide racist or daring revolutionary, but an obnoxious troll. There is no real resentment of minorities, so the white majority doesn't want to overturn society and its dominant socio-political trends of the last 50 years. They just want to break the elite which has ossified and exacerbated the worst manifestations of these trends.

Granted, this is a whole lot of speculative trend-building conjecture on Caldwell's part. Many people could read the preceding paragraphs and declare it to be a delusional inversion of a



reality wherein America has grown progressively more racist in reaction to rising black status, and Trump is the ultimate manifestation of white people's rage at losing their place at the top of the social hierarchy.

I slightly lean towards Caldwell's take, though I'd guess that both sides assign far more socio-political coherency than there really is in reality.

## Judicial Politicization

Earlier I described how the civil rights movement pioneered a practice of manufacturing or highly selecting incidents to be used for court cases to curry sympathy from the public. This strategy might sound familiar to anyone who knows the 2005 case, *Kelo v. City of New London*, or anything the Institute for Justice does. IJ isn't a civil rights group, but an organization that uses the civil rights movement's judicial strategy for libertarian policy goals. The Institute scours the country for maximally sympathetic individuals wronged by government regulations and sues on their behalf in the hopes of using the particulars of the case to sway public opinion to their side. For instance, they've defended private hair braiders and florists from obstructionist bureaucrats trying to stop dangerous maniacs from touching hair and arranging flowers without licenses.

I personally think the Institute for Justice is doing great work. But Caldwell would point to it as evidence that the entire judicial system has been hopelessly politicized. The civil rights movement moved much of the legislature's lawmaking power to the court system, established orchestrated case-building as an accepted norm, and entrenched the judicial system as a backdoor to American governance. As a result, the judicial system has been reorganized around PR battles backed by powerful special interest groups dumping money behind particular cases for their own ends. So now all sides of the political spectrum battle for control over the courts, from libertarians like the Institute for Justice to religious organizations grabbing their own anti-discrimination victories. Witness the perpetual war over appointing justices to the Supreme Court. Legal issues as big as the *legalization of abortion* have been entirely ceded by Congress to the courts, so the right and left must fight and die on the judicial hill at every opportunity. Indeed, Caldwell considered *Roe v. Wade* the final death blow to judicial impartiality.

According to Caldwell, prior to the 1960s, judges were appointed by the same criteria as sheriffs in Wild West frontier towns in movies. They were supposed to be stern, incorruptible, honorable, honest men who abided by their sworn duty of upholding the Constitution. The idea of a politician appointing a judge because of his ideology or politics was considered a soft form of corruption.

I have no idea if that's true; it sounds vastly over-idealized to me. But still, there's probably a kernel of truth to it. Do we really want the judicial system to be a weird pseudo-legislative body where lawyers backed by special interests try to convince clueless juries and politically appointed judges to make our laws?

## Financial Time Bomb

Caldwell argues that Reagan threw America into a dangerous cycle of deficit spending for the sake of maintaining the civil rights movement's social programs while cutting taxes to appease his conservative constituents. But Caldwell's causal arrow from Reconstruction 2.0 to ruining America doesn't stop there. Because how does America pay for deficit spending? By borrowing and the Federal Reserve's production of cheap credit.

The perpetual flood of cheap money hot off the Fed's brrrrring money printer fundamentally reshaped the American economy into a hyper-leveraged debt-ridden house of cards which has ushered in a new age of booms and busts that America was supposed to have left behind 100 years ago. After a mild recession in the early 90s, we had the collapse of the dot-com bubble, then the epic Great Recession, and now (as of writing this) we might be staring into another economic precipice as the corona virus grounds the global economy to a halt. To Caldwell, this can all be reframed as America handing the country's economic reigns to the cowboys of Wall Street, just so the government can have cheap credit to finance deficit spending on entitlement programs we can't afford all while building towards the greatest debt crisis in human history. The scariest part of all is that *there is no solution*. No president or Congress is going to get elected on a platform of fiscal responsibility. No one is going to slash Social Security, Medicare, and Medicaid, the three expenditure items that take up about 2/3rds of the federal budget unless they have some sort of political kamikaze death wish.

Cutting budgets is always hard for the government, but the US federal government did successfully cut a significant portion of its expenditure after the Civil War, World War I, and to a lesser extent, World War II. *But the US has never successfully cut its budget during the age of entitlement*. It doesn't matter how good or bad the economy is, or whether the US is waging a war, expenditure simply won't go down anymore. Unless the US government is some sort of economic reality-warping anomaly that can borrow evermore money forever, all these unfunded liabilities will have to be paid for eventually, either by higher taxes, inflation, or something even more extreme. And to Caldwell, this dire end will be a direct consequence of the age of entitlement.

## Coming Apart

The greatest impact of the 1960s on America is hard to explain, and Caldwell doesn't even attempt to summarize it succinctly. By my best conceptualization, Caldwell believes that the civil rights movement undermined the longstanding unifying cultural forces that tied the racially, ethnically, ideologically diverse American population together. The result is an American culture that is more polarized, divisive, and melancholic than ever before. Basically, now everyone hates each other.

One of the great things about Enlightenment-based negative individual rights is that everyone can have them without conflict. "My rights end where your rights begin," as the saying goes. But the same is not true of entitlements. If one American is entitled to a welfare check, then another American is forced to pay it. If one American is entitled to a job, then another American is forced to provide it. Thus, legal rights and civil society have become a zero-sum game. And as in all zero-sum games, people fight over finite resources. Or in this case, Americans fight over finite entitlements.

This is why America's political system has descended into polarization. There is no unifying basis or goal for American governance anymore. Protection of individual rights isn't the ultimate objective, but at best a right-wing talking point. The screaming pundits on Fox and MSNBC are symptoms, not causes. They constantly shout and insult because all their side can do is claim their piece of the entitlement pie and desperately fight off the other side. This fight extends beyond the political sphere and deep into civil society where all culture is political. All Twitter users can wield mobs or be crushed by them based on the right or wrong utterances. All Netflix tv shows make political statements with the racial composition of their casts.

This is an America addicted to entitlements: a divided, enraged, ever-fighting America. The

fight occur along racial, class, ideological, and many more lines, but most importantly everyone needs to fight. Because if you're not winning entitlements, you're paying for them.

### III.

#### An Alcoholic Aside

Without going too much into a tangent, I find the case study of substance prohibition to be an interesting parallel to Caldwell's arguments which he unfortunately doesn't address.

Various drugs had endured a wide variety of legal classifications in America throughout the 19th century, but enforcement on restrictions or bans were always lax. Americans just didn't seem to care enough about substance abuse to do anything about it.

That is until the rise of the Progressives in early 20th century America. The temperance movement aggressively campaigned against the evils of alcohol consumption, which supposedly (and not unreasonably) caused unemployment, crime, spousal/child abuse, sickness, and premature death. In 1920, the Progressives achieved an enormous and unprecedented victory by successfully outlawing the sale of alcohol in the United States. According to Caldwell's world view, this would constitute a massive violation of the fundamental principles of the Founding Fathers and individual rights.

However, *unlike* the blatantly unconstitutional Civil Rights Act and its supporting court cases (at least in Caldwell's view) the massive restriction on liberty instituted by alcohol prohibition was achieved by a *legitimate Constitutional process*. The Progressives couldn't pass a law as blatantly unconstitutional as alcohol prohibition, nor could they brute force their way through the courts without support from wild activist judges. Instead, they went through the correct process and changed the Constitution itself with the passage of the *18th Amendment* via 2/3rds majority in Congress.

Though alcohol consumption rates fell surprisingly steeply, drinking was still widespread, and 1920s America gained a culture of notable licentiousness and debauchery known as the "roaring 20s." Respect for the rule of law plummeted as speakeasies sprung up to supply the newly illicit substance. Organized crime boomed and gave birth to the American Mafia.

Many Americans considered prohibition to be a noble experiment, but with crime ballooning and drink rates creeping back to the pre-prohibition levels, public sentiment decisively turned against the policy. In 1933, alcohol prohibition was ended through Constitutional means with the passage of the *21st Amendment*. So, a minority of Americans wanted to introduce a massive change to society which would directly violate contemporary law. This minority got enough other Americans onboard with their plan to enact their dream policy through the righteous, approved legal channel ordained by the country's founding document. Americans found that they didn't like the policy, and went through the same legal channel to end the policy.

I believe this is what Caldwell wishes the civil rights movement had done. If the Civil Rights Act, or some alternative form of it, had been passed via a Constitutional Amendment, then it would have required a larger degree of support from the voting population. As a result, it would have been less radical, and wouldn't have carried the array of expansionary court cases with it. Plus if the American people decided they didn't like all or part of the Amendment, they would have a clear pathway to reversing it with another Constitutional Amendment, rather than having to slay

the hydra of a million court precedents and legislatively-imposed regulations.

I'm curious if there are any other issues or policies where the US people and government went through a similar arc. I'm not sure.

## IV.

### Priors Adjustment

As I said all the way at the start of this essay, I don't know if I buy much of the story in **Age of Entitlement**. Caldwell makes a lot of huge claims which would take a dozen more books on legal history, legal theory, and political science to thoroughly prove, and he doesn't really put much effort into demonstrating the causal chains himself.

Nevertheless, I found **Age of Entitlement** fascinating. The main effect it had on me was to lower my confidence in basic, foundational history that I learned all the way back in grade school and haven't seriously thought about since. Caldwell doesn't present a ton of evidence, but he tells a compelling story; and in retrospect, the level of evidence and storytelling is about the same as what I was taught in school prior to college.

The book also made me further reevaluate the ideological priors of my understanding of the standard historical orthodoxy. Now more than ever, I get the sense that a lot of historical analysis, especially social analysis, is done at the level of Caldwell or modern political pundits, with lots of conjecture and storytelling used to tie together distant data points with strands built from ideological preconceptions.

I'm sure Caldwell is wrong about a lot, but he's spot-on about how the 1960s era is understood and taught in America. His descriptions of textbooks and teachers closely mirrored my own experiences. It's easy to see how predominant social norms surrounding race, welfare, and the legacy of civil rights could make a historical explanation like Caldwell's completely radioactive in any American school. Even considering such ideas would bring down a deluge of accusations of racism, sexism, homophobia, etc.

Maybe that's what all history is: weaving together data points with ideology. Whatever ideology has dominance in a particular setting gets its story told.

# The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia by James C. Scott (2009)

I decided to read this because Scott's book "Seeing Like A State" is one of the best books I've ever read. If this sounds interesting, I highly recommend reading that book first.

Scott does not write history books that tell you what happened. Scott writes history books that tell you what broad patterns to look for in history.

The key idea in Scott's work is legibility. Organically grown human societies are enormously complex. There are many local variations, grown from the preferences of many individuals. Although this makes sense to the people who live in the society, it does not to outsiders. Large institutions, especially the state (meaning government - not the US meaning), cannot function in this environment. In order to do anything, they first have to make society legible to a central administrator. So they bulldoze the organically grown society. Many aspects of our lives were created by the state forcibly replacing the organic society with something more amenable to the state - cities planned on a grid, monoculture agriculture, unambiguous property rights, standardized measurement systems, national languages with dictionaries & grammar rules, and even unique unchanging names. But, in order for the state to thrive, the individuals also have to thrive. When administrators push to "rationalize" society too much or too quickly, the results are catastrophic.

---

"The Art of Not Being Governed" is Scott's discussion of the area of the world where he has the most expertise: peoples living above 300m (about 1000ft) in mainland Southeast Asia, southern China, and northeastern India.

Two features of the hills of Southeast Asia are immediately obvious:

(1) It has extremely difficult terrain: steep mountains and dense forest.

(2) It has incredible cultural and linguistic diversity. Even a map like the one here underestimates the complexity because many villages (and people) are multilingual:

[http://sabotin.ung.si/~jezik/slike\\_arhiv/jezikovni\\_nacrti/Huffman-SouthEastAsia\\_Langs-wlms-150.pdf](http://sabotin.ung.si/~jezik/slike_arhiv/jezikovni_nacrti/Huffman-SouthEastAsia_Langs-wlms-150.pdf)

The history of the region is typically told from the perspective of the civilized valley states, which have wet rice agriculture (padi/paddy), hierarchical social structure, a literate elite, and orthodox religion. The uncivilized hill people, who have transient agriculture (swidden/slash-and-burn) or foraging, sometimes egalitarian social structure, no writing, and heretical charismatic religion, are seen as a relic of a more primitive time that is slowly fading away as the people learn the benefits of civilization.

This perspective is wrong. The hill peoples are not primitives who aren't yet civilized. Most of the peoples in the hills previously lived in the valleys and moved to the hills fleeing state violence. They have been in cultural and economic contact with the valleys for thousands of years. If these peoples aren't civilized, it's because they actively chose not to be, not because they never learned.

The hills of Southeast Asia should instead be thought of as a shatter zone. The incredible linguistic diversity is the result of wave after wave of peoples fleeing state violence. They mostly came from what is now southern China, but also from the padi states of Vietnam, Thailand, and Burma. During good times, some people would be drawn from the hills to join the valley people (physically, linguistically, culturally, even ethnically). But that state can be incredibly violent: high taxes, corvée labor, epidemics (from concentrated population), slave raiding, civil war, large armies that live off the land, etc. During most of Southeast Asia's history, more people choose to move to the hills than to move to the valleys. States would often collapse when poor governance led to most of the population running away. Even the states that did survive waged frequent wars with the intent to capture slaves to work the wet rice fields (until they ran away).

Other shatter zones exist. They tend to be mountains, dense forests, or swamps at the margins of major states which have incredible cultural diversity. The central mountains of India, Afghanistan, the Caucasus Mountains, and the Balkans are all clear examples. Switzerland and the Low Countries might be shatter zones, although most people in Western Europe fleeing the state have moved overseas for the last 500 years. How many of the native societies in the Americas were original vs. post-epidemic shatter zones is widely debated, but the Seminole in the Everglades and the runaway slave colony in the Great Dismal Swamp are more recent American examples. I don't know enough about African history to give examples there. Nomadic societies, whether on horseback or in boats, share many features with shatter zones. They incorporate a wide variety of people fleeing states, but tend to form larger, shifting confederations instead of many small, diverse communities. Although these societies may look primitive for the perspective of civilized people, they postdate states. Steppe nomads require domesticated horses and swiddening in hardwood forests requires iron axes, both of which were developed after the first states.

Several aspects of the cultures of the hills of Southeast Asia are particularly illustrative: farming practices, political/social structures, ethnicity, writing, and religion.

## Farming

In Southeast Asia, the factor limiting food production was labor, not land. This has only started to change in the last few decades. This explains why it was so easy to run to the hills and why collecting (often forced) labor was so important to states.

Individuals care about efficiency per hour worked. States care about efficiency per land area.

Efficiency per land doesn't matter for the people who live among the fields. It matters for people who try to tax food and gather it to the capital. Before the introduction of railroads and motorized vehicles, moving food overland was prohibitively difficult. Armies had to march immediately after the harvest or they wouldn't be able to provide for themselves.

Wet rice is the perfect crop for the state. Not only is it extremely high density, it also is extremely legible. The rice is all planted and harvested at the same time, so it is easy to assess and later collect the appropriate taxes for each padi. Although wet rice is easiest to grow in flat, wet valleys, with enough labor, the hillsides can be carved into terraces and filled with stream water. Wet rice is almost always found in states or among peoples who aspire to form states.

Subsistence in the hills is much more varied. Some people practice fixed field agriculture. Some people practice transient swidden agriculture. Some people are foragers in the forest. People switch between these practices, or use a combination of them.

Many of the crops grown are from the New World - especially potatoes, maize, and cassava. These not only do better at higher altitudes than most Old World crops, they are also much less legible than rice. Root crops, in particular, can be left in the ground instead of harvesting them all at once and placing them in a storehouse. This makes it much harder to collect taxes - or for raiders to steal your food. Crops that require little care can even be abandoned if you have to go hide in the forest, and still have food when you get back.

What crops are planted and what farming or foraging practices are used allow people to determine how legible their food source is. If the nearby states are attractive, they might switch to more settled agriculture, or even move into the padi core. If the nearby states are dangerous, they might abandon all farming except for a few hidden fields of potatoes and cassava and get most of their food from foraging.

Regardless of their agricultural choice, hill people remained connected to the valleys through direct or indirect trade. The hill people needed iron tools from the valley forges. Much of the

wealth of the valley states came from international trade in luxury goods that could only be found in the hills.

## Political structure

States are inherently hierarchical. There is a central authority which monopolizes force to impose its will. This is especially true in the Indian tradition of Theravada Buddhism, with an ideal universal ruler in the center of the wheel, and in the Chinese tradition of Confucianism, with its emphasis on proper action in hierarchical relationships. Liberal democratic government was unknown.

States prefer to interact with other hierarchical organizations. If a tribe has a single chief, then it is clear who the state should negotiate with. If there is no hierarchy, then there's no one with the authority to sign a treaty detailing their relationship to the state. Egalitarian social structures not only prevent the emergence of a state within your own society, they also make it more difficult for neighboring states to exert their influence.

States, both traditional and colonial, actively encouraged the peoples surrounding them to organize into tribes with an unambiguous chief. Some peoples did, becoming miniature kingdoms in the hills. Each one would be allied with one of the neighboring valley states. Without the economic concentration of padi agriculture, political concentration is difficult to maintain without outside support. Others actively resisted this and maintained traditions of killing anyone who aspired to become chief.

In the absence of a political hierarchy, social status was often determined by competitive feasting. Whenever someone had an excess of food (from a successful hunt, a bountiful harvest, new trade goods, etc), they would invite everyone in the village to a feast. A grand feast would be remembered, and increase your social standing in the group. These feasts had the effect of exchanging unequal material wealth for social status. A sure sign of increasing hierarchy is when someone tries to restrict the right to hold feasts to only some members of the group.

Much like their agricultural practices, hill peoples could choose their social organization based on how much state influence they wanted.

## Ethnicity



One of the goals of colonial governments in Southeast Asia was to take a census of the hill peoples and categorize them based on their ethnicity. This proved to be completely impossible.

Ethnic groups are supposed to be (mostly) genetically and culturally distinct groups. Even if the peoples do not stay as distinct as they ought to be, when there is cultural variation, you'd think that you could find some way to categorize them.

There are various ways to distinguish between ethnicities, for example by dress or food or language. Even if you do choose how to divide up the different styles of dress, these lines do not follow the divisions in food, or language.

But this is not the biggest problem for the census taker. Most people in the hills are multilingual. And not only can they speak multiple languages, they know how to function in multiple different cultures.

The individual people have command of multiple cultural repertoires that they can use in different contexts. These contexts might be for a short time scale, like who I'm talking to or what sort of event is happening, or they can be for a long time scale, like whether I'm going to spend the next few years swiddening in the mountains or farming rice in a valley state.

This completely undermines what we typically mean by ethnicity. Ethnicity is something you are born into. It is not something that you can change whenever you feel like it.

Although these behaviors make sense to the people doing them, they are extremely confusing to the census taker. How do you categorize someone with this cultural flexibility. The resulting surveys are comical: small villages which contain almost as many ethnic groups living side-by-side and interacting daily as they have households; entire societies that vanish and later reappear without any loss of life; and, of course, the people who purposely make life more difficult for the census takers, because they realize that the goal is to make it possible for the state to exert its power over you.

## Writing

This chapter of Scott's book is more speculative, and relies on ancient local legends and personal speculation more than lived experience. But his ideas are interesting, even if they are not true in a particular situation.

The valley states see the hill peoples as pre-literate. They have not yet learned to write. The legends of the hill peoples typically describe themselves as post-literate. They used to know how to write, but lost it due to treachery, trickery, or neglect, typically involving a dramatic story

with supernatural elements.

Writing, and history in general, are also ways to control how legible your society is. Even in the most civilized states, literacy was exclusive to the elite. Writing helped to create and maintain a hierarchy.

Writing also institutionalizes society in another way. A written text provides a long-lasting source of authority to check claims against. You cannot reinvent your history if there is a text to check it against. Whoever controls the text has epistemological influence over society. Losing the texts, whether because they were eaten by pigs, or because a trickster stole them, or because a neighboring king took them and executed every literate person in the tribe, makes the society more egalitarian.

Even in the absence of writing, some hill peoples maintained extensive oral traditions. These could be long and detailed: the story of a murder that sounds like a police investigation or genealogies going back 10 generations. This provided an alternate form of maintaining history.

Other peoples kept no history at all. They kept no stories of great deeds (unless it involved murdering an aspiring chief). They refused to remember genealogies back beyond the people they knew personally. They even would not think of themselves as a people - the names we use to refer to them are entirely given by outsiders.

Illiteracy is one of the main things that the valley states use to stigmatize the hill peoples. Many of the hill peoples agree and see it as a source of shame. It can also be seen as another source of illegibility. People without history are as flexible as possible. They can change subsistence practices, alliances, and even their ethnicity more easily than people who are constrained by an existing text.

## Religion

The padi states tend to practice an orthodox version of one of the major world religions. Which religion varies between states: Confucianism in Vietnam, Hinduism and then Theravada Buddhism in most of the region, and Islam in Malaysia and Indonesia. Regardless of which religion is practiced, it is always an orthodox version. The religion is a hierarchical institution with written commandments and an exclusive class of priests/scholars. A foreign source for the religion also helps to invoke the legitimacy found in better established states.

The religions of the hills tend to be different from the neighboring padi states. Animism and Buddhism are common in the hills throughout the region, along with Hinduism in the hills of Indonesia and Islam in the southern Philippines. Christian (especially Baptist) missionaries have

found most of their success in the hills - except for in the Philippines, where Catholicism became the valley religion.

The religions in the hills are never orthodox. They blend multiple religious traditions and local folklore. There is rarely any established hierarchy. Buddhist monks in the hills tend to be self-appointed hermits, rather than members of a monastic order.

Hill religions also have strong messianic / prophetic tendencies. Belief in a future messiah is common among peoples who have been marginalized, oppressed, and driven from place to place. Prophets, who gain followers by their personal charisma, provide a way to organize people who were notoriously opposed to institutions. These prophets might trigger revolutions (appealing to both hill people and the poor in the valleys), lead groups to settle new territory, promote new customs, and establish new legends that would be incorporated in all later religions. Following a new prophet allows for the immediate creation of a new culture and society in response to the immediate circumstances.

Christian missionaries found particular success in the hills. Scott discusses several possible reasons why. Christianity is a messianic tradition. Looking forward to the Second Coming of Christ resonates with previous traditions of a future ideal king. The most successful missionaries in the hills were typically Baptists, who promote local autonomy over institutional authority. Not only is this consistent with the illegibility of the hills, it also makes it easier for individual churches to incorporate earlier traditions. Christianity allows the hill people an alternative route to modernity, escaping the stigma from the valleys, without becoming subject to the padi states. In particular, missionaries would bring or create alphabets, bringing literacy to a people (not just an elite) who had previously lost their letters.

{ This got me thinking about the relationship between various religions and how widespread literacy is. In particular, Abrahamic religions tend to promote mass literacy. The Torah commands the Levites to read the entire law to the people every 7 years (Deut 31:9-13) and encourages at least a little interaction with text among all the people (Deut 6:6-9). This has developed into the frequent practice of Torah reading (in Hebrew), which continues in modern Judaism. Having the people read the scriptures continues to be a theme in the New Testament - which encourages mass literacy (Luke 24:35, Romans 15:4, repeatedly in the letters to Timothy). Medieval Catholic use of the Latin, instead of vernacular, Bible is a deviation from this pattern which has been corrected. The Book of Mormon is even more explicit: a prophet tells poor people who have been kicked out of their congregation to search the scriptures themselves (Alma 33:2). In Islam, Mohammad's first revelation involves Gabriel berating him for not knowing how to read (Quran 96).

I don't know enough about the history of other religions to be able to look at their relationship to mass literacy. I would suspect that at least some of them traditionally encourage literacy for the elites, like the Hindu brahmin or the Confucian scholars or the Roman augurs, and discourage literacy for everyone else. This would fit with the trend Scott describes of institutional religions being used to protect the status of the elites relative to the commoners. To be clear, I don't

know about these particular examples. If anyone is familiar enough with pre-modern non-Abrahamic religious practices to comment on this, I would love to hear it.

If my suspicion were true, then we would expect pre-modern societies with Abrahamic religions to have significantly higher literacy rates than other pre-modern societies. Scott unknowingly provides evidence for this claim by saying that the places with the highest literacy in pre-colonial Southeast Asia were Indonesia and the southern Philippines, which were the areas where Islam was established. It would be interesting if mass literacy, one of the hallmarks of modernity, actually dates back to the Torah. }

Two quotes from this chapter, in particular, stood out to me:

"It was the good fortune of the Baptist missionaries to have brought the Bible to a people who had long believed in messiahs. It was their mistake to imagine that the Baptist messiah was the last messiah the Karens, in their impatience, would be needing."

"The frequency of Lahu prophetic movements allows us to identify something of a 'career trajectory,' even for an activity so decidedly unroutine as becoming a god-man."

---

Southeast Asia is rugged terrain next to the oldest and largest state building project: China. As the Han Chinese expanded southwards from the North China Plain, wave after wave of ethnic groups, failed rebels, and oppressed farmers fled from the padi state to the hills, despite the efforts of the state to contain them. Smaller, closer padi states, like Vietnam, Thailand, and Burma, also contributed. These refugees created a human landscape of remarkable linguistic and cultural diversity. It has only been recently that modern transportation and communication technology has allowed the state to extend its influence into these regions, and that is still very limited in some areas.

While most aspects of life in the padi state are designed (by the state) to be as legible as possible to a central administrator, life high in the hills tends to be as illegible as possible, with a continuum of options in between. People can choose between the wet rice of the padi state, dry fixed field agriculture, transient swiddening agriculture, and foraging in the forest. People can choose hierarchical, inherited social structures or more egalitarian systems where social status is based on competitive feasting. People can choose from or mix multiple world religions and local traditions. All of these choices can be made either as a group or individually. Most individuals are multilingual and can function equally well in multiple cultures. They can easily choose which hill society they want to live in, or whether to move to a valley state, changing their ethnicity at will.

The societies of the hills are not primitive. They were created by intentional choices of people

who were familiar with and rejected the state.

# The Book Of Why by Judea Pearl

[Here](#) is a strange coincidental story from a couple of days ago mentioning the author Judea Pearl. Unfortunately he seems most famous for some shitty terrorist act which killed his son. I mention it only to emphasize the important fact that weird coincidences do happen. But let's quickly set that aside and remind ourselves who the author, Judea Pearl, really is.

[Wikipedia tells me](#) that Pearl won the Turing Award: "For fundamental contributions to artificial intelligence through the development of a calculus for probabilistic and causal reasoning."

"The Book Of Why: The New Science Of Cause And Effect" is Pearl's best attempt to communicate what that means to people like me. I get the feeling that communicating to regular people is not Pearl's primary strength (observe there is a co-writer, Dana MacKenzie).

What is the point of the book? What problem does it try to solve? Well, it turns out that it's quite topical. Take the following example. An excellent author I've [previously reviewed](#), Ed Yong, [writes in the Atlantic](#):

The French studies that first suggested that [hydroxychloroquine] could treat COVID-19 were severely flawed, abandoning standard elements of solid science like randomly assigning patients to receive treatments or placebos, or including a control group to confirm if the drug offers benefits above normal medical care. The lead scientist behind those studies has railed against the "dictatorship of the methodologists," as if randomization or controls were inconveniences that one should rebel against, rather than the backbone of effective medicine.

For the last century or so, statistics has indeed assumed that if you really want to understand if a drug *causes* a beneficial effect, you needed to do a [randomized controlled trial](#). Statistics textbooks do not admit any other way. But what if there was a way? Judea Pearl thinks that sometimes there is. That'd be pretty important right about now.

The real core of the book is causality. Did the drug *cause* the recovery? To understand why this is a complicated question, consider this wonderful example of bad causality - a crystal health index. [This article](#) basically constructs a crystal health index where medical outcomes are tied to things like exercise, diet, and proximity to crystals. Clearly this crystal index will correlate to good health, but is that because crystals cause good health? No, it is a flawed metric by design to illustrate the problems with such things.

Here's another critique of relying too heavily on randomized controlled trials. [Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials](#). Go ahead and read the abstract — it's pretty funny. But the point is

that sometimes the ideas behind RCTs are overkill. Pearl would like us to think of these problems in a different way and has some interesting insight.

Here's another example I came across. Check out [the interesting history of bone marrow transplants](#). It describes when this painful procedure was in fashion, even though it is now generally regarded as worse than useless. This quote sums up a conundrum both for treatments that work great and bogus ones.

So the procedure wasn't just **untested**, it also **couldn't be easily tested**, because what patient enrolls in a clinical trial where you get assigned at random to a control group, where you **wouldn't** get a live-saving procedure? Do you want to be the dead data point which allows someone else to be the live one?

So how did we get here? When guys like Pascal, Cardano, Fermat, Huygens, Halley, deMoivre, Bernoulli, LaPlace, and Gauss worked on probabilistic thinking, they really were just trying to make some sense out of card games for their rich gambling patrons. Their limited success was cool — making sense out of more complicated things seemed promising.

Ironically, the need for a theory of causation began to surface at the same time that statistics came into being. In fact, modern statistics hatched from the same causal questions that Galton and Pearson asked about heredity and their ingenious attempts to answer them using cross-generational data. Unfortunately, they failed in this endeavor, and rather than pause to ask why, they declared those questions off limits and turned to developing a thriving, causality-free enterprise called statistics.

— p4

Francis Galton (Darwin's cousin) invented the linear regression. His student Pearson invented the (Pearson) correlation coefficient. Pearl believes these guys set up "scientific" thinking — probably better called "natural philosophy" for this discussion — to only accept randomized controlled trials. Rightly or wrongly, this effectively shut down other possible ways of inferring causality. Pearl believes this was an unfortunate error.

[Traditional statistics practitioners] believe that the legitimacy of causal inference lies in replicating a randomized experiment as closely as possible, on the assumption that this is the only route to the scientific truth. I believe that there may be other routes, which derive their legitimacy from a combination of data and established (or assumed) scientific knowledge.

— p334

Pearl tries to more accurately identify what the essential magic is with RCT.

"...the principle objective of an RCT [randomized controlled trial] is to eliminate confounding..."

— p150

Confounding is "...the discrepancy between what we want to assess (the causal effect) and what we actually do assess using statistical methods." Confounding is a discrepancy between these terms.

$$P(Y|X) \neq P(Y|\text{do}(X))$$

The first term is the probability of Y given X. E.g. the probability of dying given that a patient was in icy water for 10 minutes. The second term was more problematic to me. It seems to mean: the probability of Y given *action is taken to ensure* X. So, e.g. the probability of dying given that someone was thrown overboard into icy water for 10 minutes. Subtle difference — at least statistically. Filed under [probabilistic causation](#), Wikipedia says: "**do(X)** stands for an external intervention that compels the truth of X."

Mmmm Hmmm. Okaaay. Is this notion and notation important? Pearl emphatically tells us yes!

You might say it's obvious or common sense, but generations of scientists have struggled to articulate that common sense formally, and a robot cannot rely on our common sense when asked to act properly.

— p151

I love his computer science spirit there but unfortunately the book was filled with hand waving "math" and no real code that "a robot could rely on". Sure, I guess the translation is theoretically possible but not by me.

I liked this little plug for his pet do-operator.

Lacking a [causal] diagram or a do-operator, five generations of statisticians and health scientists had to struggle with surrogates, none of which were satisfactory. Considering that the drugs in your medicine cabinet may have been developed on the basis of a dubious definition of "confounders," you should be somewhat concerned.

— p152

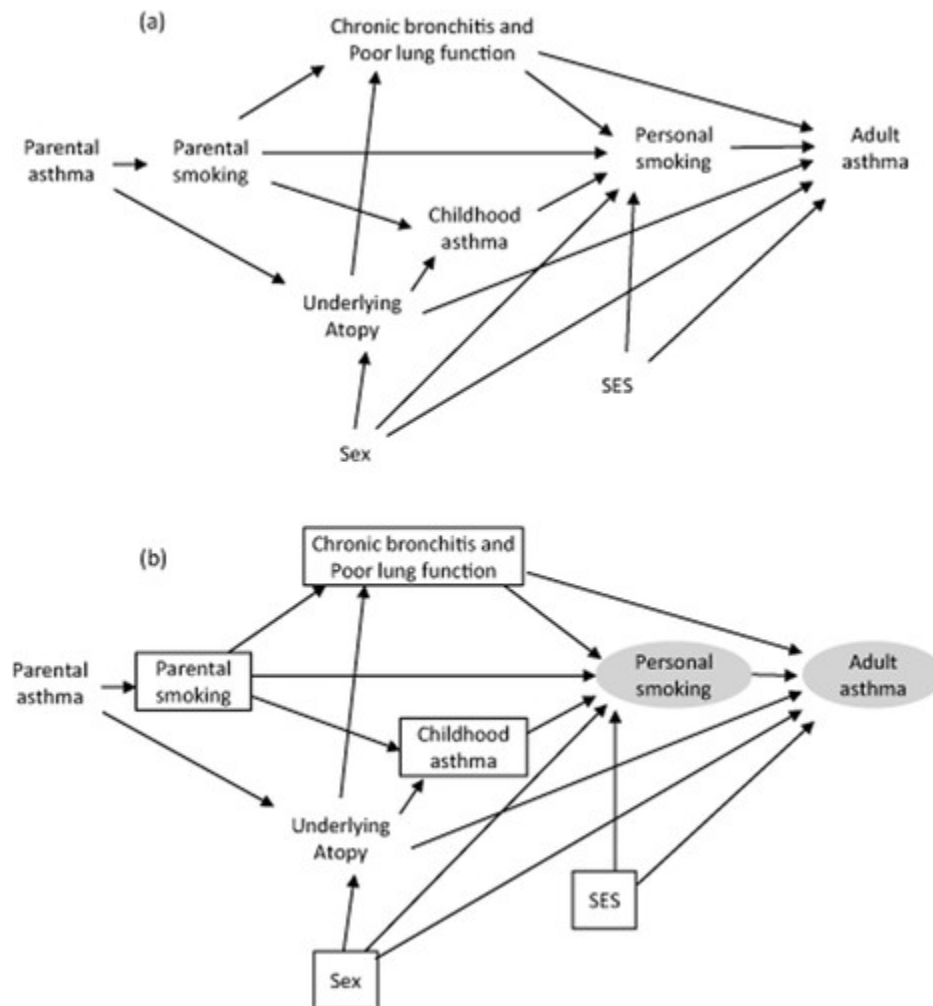
I can tell you [a lot of other reasons](#) you should be concerned too!

Ok, so there's this **do** operator which makes the flaky weird world of statistical notation just a little bit flakier and weirder.

To me, what was more immediately comprehensible was how Pearl invited some graph theory into the party. For all my irritation with boffins hiding behind abstruse mathematical glossolalia, graph theory is actually totally wholesome.



Pearl proposes some kind of causal graph is a useful prop. Here's one that I found (not in the book) which shows a lot of stuff going on.

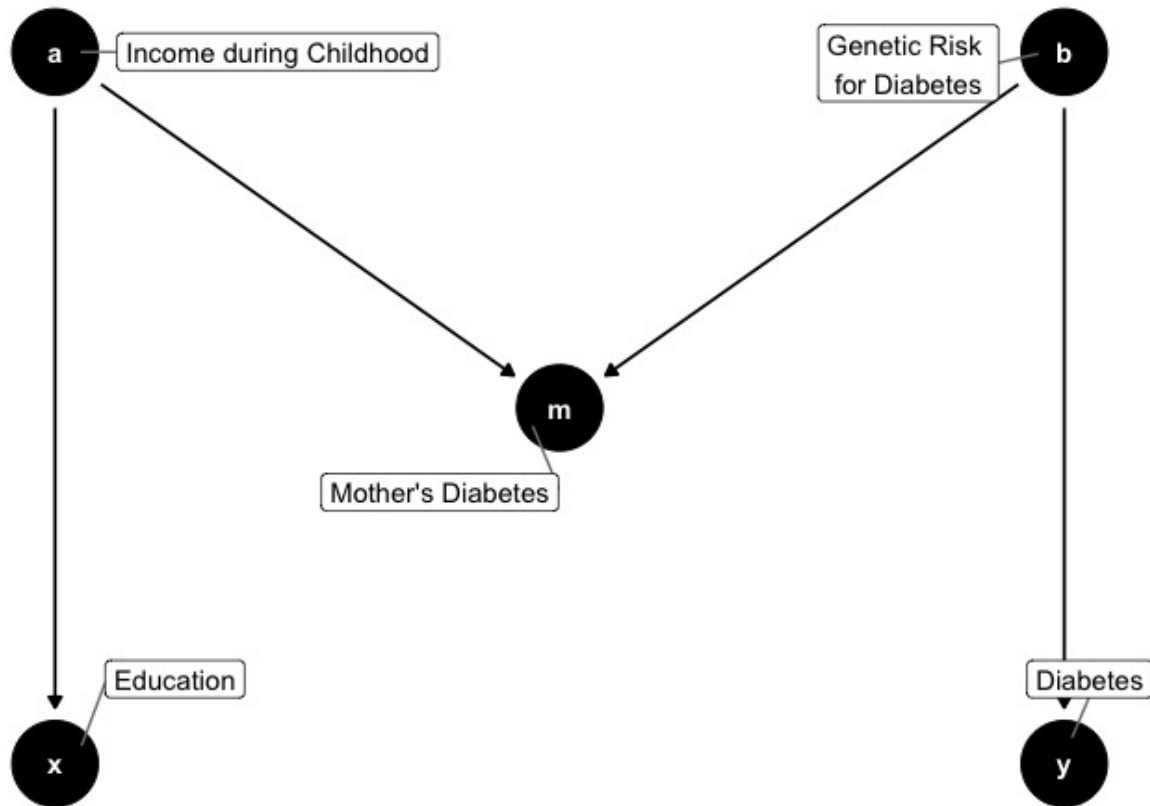


I believe that the arrowhead points to the thing that the arrow's tail causes. Pearl believes there are some fundamental arrangements of causal graphs and that many logical errors can be rooted out by thinking in this way.

Here is how the graphs relate to the **do** operator, and how all that relates to something like randomized controlled trials.

The do-operator erases all the arrows that come into X, and in this way it prevents any information about X from flowing in the noncausal direction. Randomization has the same effect.

Here is one example of a motif he highlights (on p161). It is an M-bias arrangement.



Apparently all statisticians before 1986 would consider M a confounder because it is correlated to X and Y, however there is in fact no causal path from X to Y.

M-bias puts a finger on what is wrong with the traditional approach. It is incorrect to call a variable, like M, a confounder merely because it is associated with both X and Y. To reiterate, X and Y are unconfounded if we do not control for M. M only becomes a confounder when you control for it!

[This resource](#) talks more about this in more detail. And a lot of examples of causal diagrams can be found [here](#).

It turns out that this kind of analysis is very important in many different applications. Look at this example where it is pointed out (p261) that the concepts of "necessary", "sufficient", and "necessary and sufficient" are all related to causality.

Using these words, a climate scientist can say, "There is a 90 percent probability that man-made climate change was a necessary cause of this heat wave," or "There is an 80 percent probability that climate change will be sufficient to produce a heat wave this

strong at least once every 50 years."... Either of these statements is more informative than shrugging our shoulders and saying nothing about the causes of individual weather events.

What got me interested in Pearl is hearing his work mentioned with the idea of "counterfactuals". Counterfactual events are events that could have happened but did not happen. It turns out that for causal reasoning with imperfect knowledge, thinking carefully about counterfactuals can be useful. The details of exactly how are still not quite clear to me. For example, no mentally healthy person could claim to properly understand [this kind of writing on the topic](#). But there are important real world considerations.

Pearl is actually kind of down on AI research — no doubt because his academic flavor of it didn't step into a pile of high visibility commercial success.

[AI researchers] aimed to build robots that could communicate with humans about alternate scenarios, credit and blame, responsibility and regret. These are all counterfactual notions that AI researchers had to mechanize before they had the slightest chance of achieving what they call "strong AI"...

— p269

But he's 100% right about that. He further critiques modern AI miracles.

In technical terms, machine-learning methods today provide us with an efficient way of going from finite sample estimates to probability distributions, and we still need to get from distributions to cause-effect relations.

— p262

This stuff is complicated. The primary motivation to study causality is predicting the future — how can we avoid the bad situation and cause the good one *in the future*? Predicting the future is obviously complicated, but even talking about how well you did at predicting the future is very tricky. So all of these kinds of discussions are problematic.

I was impressed that Pearl did not accept the state of the art in statistical thinking in 1900 as the last word on the topic. That correlates perfectly with my attitude! I was impressed that he had some plausible ways forward. He is trying to tie in some thinking from other disciplines. I like that approach.

What was disappointing was that none of this is new (he's already won the Turing Award for it) and I was hoping this book would help me make actionable progress using his strategies. Maybe that will happen for others, but not for me. The best I can say is that I'm aware of this kind of thing now and can follow along with newfangled causal reasoning.

I was also disappointed that there was no explicit algorithms, code or pseudo code examples. That would have helped a lot. It's easy to find mumbo jumbo on complex math and philosophy

topics. Finding some that makes a robot do the right thing injects instant credibility into the conversation.

My suggestion for someone who attempts to write about this topic to a general audience is to go to basics — a gambling game. Can a practitioner of these causal diagram methods invent a gambling game that this kind of analysis is necessary and sufficient to reliably win? That is what casinos do with (and up to the limit of) classic probability calculus.

Maybe when you read the book, you'll understand things much better than I did. Then, if you think you have a slight handle on causality, I invite you to investigate a much more complex phenomenon highlighted most visibly by small human children — [pretend causality](#)! Enjoy!

# The Burnout Society by Byung-Chul Han

Rarely does one read books that change their—already formed—views towards society. Maybe even more of a rare event is when a book makes one doubt their life philosophies and deepest goals.

[The Burnout Society](#) by Hegelian philosopher [Byung-Chul Han](#) is short yet genius. It is composed of eight chapters. Throughout these chapters the author distills an analysis into the 21st century global society that shatters the social imaginary by being both astoundingly radical and incredibly insightful.

## I.

Byung-Chul starts—in the first chapter—with a core argument of the book's thesis:

Neurological illnesses such as depression, attention deficit hyperactivity disorder (ADHD), borderline personality disorder (BPD), and burnout syndrome [...] do not follow from the negativity of what is immunologically foreign, but from an excess of positivity.

The author defines the era of the past century as an immunological era, which has resulted to our current era being deprived of any negativity as it has been successfully purged.

The epoch sought to distinguish clearly between inside and outside, friend and foe, self and other. The Cold War also followed an immunological pattern. [...] Even if [something] has no hostile intentions, even if it poses no danger, it is eliminated on the basis of its Otherness.

## II.

Continuing in the second chapter, Byung-Chul talks about today's society not being [Foucault's disciplinary society](#), but rather an achievement society.

In contrast to a disciplinary society, which is a society of negativity that prohibits (*don't/should*), an achievement society is one that positively encourages (*can*). Still, Byung-Chul says, the achievement-subject remains disciplined.

Clearly, the drive to maximize production inhabits the social unconscious.

With this, the author touches on the edges of Cornelius Castoriadis' concept of the [social imaginary](#)). Being achievement-subjects, we are much more productive. Discipline-subjects work because they *shouldn't* be slackers; achievement-subjects work because they *can* become millionaires.

In this way, the achievement-subject being exploits itself. Voluntarily, it becomes both the abuser and the abused.

Disciplinary society's [...] negativity produces madmen and criminals. In contrast, achievement society creates depressives and losers.

### III.

In the third chapter, Byung-Chul talks about boredom. In praise of idleness, like [Bertrand said](#)—yet not quite in the same sense.

The aforementioned excessive positivity of our society leads to excessive stimuli and information. As a result, a new form of attention develops: *hyperattention*. Characterised by rapid diffusion of information and fast change of focus, hyperattention means that our tolerance for boredom abates.

But, boredom is quite important. The author claims it to be fundamental for the spark of creativity to arise.

To handle this excess of stimuli we employ multitasking. Byung-Chul remarks that this is a step backwards in terms of cultural quality. Multitasking is common to animals who while eating must be careful not to be eaten. Deep focus instead is the intellectual technique to which we owe humanity's cultural achievements.

If sleep represents the high point of bodily relaxation, deep boredom is the peak of mental relaxation.

### IV.

The fourth chapter of the book is titled *Vita Activa*. Hannah Arendt described this concept in her book *The Human Condition*, in which she wanted to redeem it against *vita contemplativa*. According to Hannah, *vita activa* is being misinterpreted as just restlessness, whereas it can actually be heroic action.

She says modern society forces humans into *animal laborans*; a concept that designates humanity's defining characteristic as labour, rather than thinking. As *animal laborans*, there is no *vita activa*—only passivity.

Byung-Chul states that *animal laborans* are no more, though. An individual out of Hannah's model is supposed to relinquish their individuality (ego) and merge in the "over-all life process of the species". Instead, in today's achievement society ego is *cultivated* by labour.

In the second half of the chapter, the author speaks about the fleetingness of modern life. Hyperattention (chapter 3) is our society's reaction to this lack of being, as he calls it. Everyone carries a labour camp inside them, in which, he says, "one is simultaneously prisoner and guard, victim and perpetrator".

Even religions, as thanatotechnics that would remove the fear of death and produce a feeling of duration, have run their course.

## V.

Chapter five is where Byung-Chul talks about the power to say no, i.e. the power to express negativity.

According to Hegel, negativity is precisely what keeps existence [Dasein] alive.

If *positive potency* is the power to do something and *impotency* is the exact opposite, that is the inability to act, then there is also *negative potency* which is the power to not do something.

Arguments from the first and the third chapter now connect: we lack negative potency as we have purged ourselves of all negativity. Lacking the ability to say no to something (say no as in block it and not perceive it), the result is an excess of stimuli: hyperattention.

The author also touches on Zen meditation here, in which the objective to free oneself from all distractions requires activity.

## VI.

Chapter six is about [\*Bartleby, the Scrivener\*](#), a short story by Herman Melville. Bartleby, who's an above-average skilled scribe, starts responding to all requests with "I would prefer not to".

Byung-Chul analyses how the society portrayed in the story is a disciplinary society. Further, he notices a lack of excess positivity, at which point he disagrees with an analysis of the story by [\*Giorgio Agamben\*](#). According to Byung-Chul, Bartleby is not the

*absolute potency*. Him being an employee of the Dead Letter Office, also, depicts a negativity that contradicts Agamben's ontotheological interpretation.

Byung-Chul concludes that the story is not really about de-creation, but rather about exhaustion.

The exclamation that ends the tale is both a lament and an indictment: "Ah Bartleby! Ah humanity!"

## VII.

The penultimate chapter is called *The Society of Tiredness*. Bilingual reviewers say this chapter title is also a better translation of the original German book title *Müdigkeitsgesellschaft*.

Here, the author draws a lot from Peter Handke's [Essay on Tiredness](<https://us.macmillan.com/books/9781250767257>). Tiredness in our achievement society is solitary tiredness. This is in contrast to another kind; the tiredness that trusts in the world. While the trusting tiredness dissolves ego and makes room for an [Otherness]([https://en.wikipedia.org/wiki/Other\\_\(philosophy\)](https://en.wikipedia.org/wiki/Other_(philosophy))), solitary tiredness results in an exhausted ego.

[Handke] calls it "we-tiredness". I am not tired "of you," as he puts it, but rather I am tired "with you"

## VIII.

In the final chapter, Byung-Chul claims that today's achievement society is much different than the disciplinary society that Freud developed psychoanalysis for. So much different that he believes that a person's psyche is, also, much different and thus psychoanalysis is irrelevant for the achievement-subject.

In our society of achievement, there is less and less negativity and prohibition. There is instead more and more freedom. An achievement-subject does not work out of duty (defined by others; targeted at others), but rather out of pleasure (defined by itself; targeted at itself).

This absence of relation to the Other is problematic especially in terms of gratification. The achievement-subject cannot draw satisfaction from its work because there's no one to commend it. Hence, the achievement-subject feels compelled to do more and more.

This combination of lack of satisfaction and self-exploitation (chapter 2) results in burnout. As the achievement-subject is subject to no one but itself, it *liberates* itself into



a *project*. It is liberation, because the achievement-subject does that in the name of freedom, of being free to do *that* which pleases it. And it is a project—and not a subject—in the etymological sense, as it is not subjugated. It (the subject itself) needs to be something positive, after all.

In this way, though, the violence does not disappear. The violence now originates from itself. As we said in chapter 2, this is now a self-exploitation. It is represented as a freedom, yet it is still a (self-) coercion.

This result makes the capitalist system much more efficient as self-exploitation includes the illusion of freedom. Freedom and violence now coincide.

In the final pages, Byung-Chul examines Agamben's [\*homo sacer\*](#). He declares:

[The homines sacri of achievement society] cannot be killed at all. Their life equals that of the undead. They are too alive to die, and too dead to live.

## Epilogue

When I started reading this book I thought the author was just playing with words. How can immunology describe an era and somehow this metaphor to result in mental conditions?

And how does all this reconcile with current psychology? My guess would be: not well. The obvious argument that what the book describes is just a different—i.e. a philosophical—approach probably satisfies neither philosophers nor psychologists.

Yet, even though the first chapter felt silly, as I continued reading, everything made sense. Somehow, it made total sense. I remain cautious of its thesis but at the same time I feel my views towards the world have overturned. In an attempt to either verify or disprove its arguments I find myself trying to examine things through Byung-Chul's lens now.

In conclusion, I'm quite astounded by *The Burnout Society*. So much that I wonder—how is to exist before having read it?

# The Call of Cthulhu by H. P. Lovecraft

## *Mysteries of the Unknown and the Unknowable*

The oldest and strongest emotion of mankind is fear, and the oldest and strongest kind of fear is fear of the unknown.<sup>1</sup>

– H. P. Lovecraft

The first thing you think of when you hear the word **mystery** is probably the sort you find in detective novels or police procedurals. It's the kind of thing that Arthur Conan Doyle and Raymond Chandler were concerned with. Then, as you delve into the deeper recesses of your mind, you may think of religious or transcendental mysticism. Now it seems like the kind of thing that the Gnostics were concerned with. And the word's root is religious in nature, descending as it does from μυστήριον (*mustérion*), from μύστης (*mústēs*), "initiated one", from μύω (*muéo*), "I initiate", from μύω (*múō*), "I shut".<sup>2</sup> But there's nothing godly about it, nor anything furtive, nor cryptic.

This text is a commentary on H. P. Lovecraft's story *The Call of Cthulhu*. And I will get to that one soon enough, will describe it, will exalt it, will examine it. But first we must do a little groundwork, by making a distinction between two kinds of mysteries.

## A Distinction in Mystery

Consider the following examples:

- In the year 1327, a renowned friar and his novice arrive at a monastery in Northern Italy. The abbot there tells them that, not long before, one of the young monks, a talented illuminator who hailed from Otranto, was found dead by goatherds below the monastery tower escarpment at dawn. But all the windows of the tower building were closed and, though it had been raining, and though the wind blew against the windows, there was no puddle of water below any of them.<sup>3</sup>
- The observable universe is 93 billion light-years in diameter and contains trillions of galaxies, each of which galaxy holds billions of stars, around one of which orbit eight huge rocks. One of those rocks is Earth, which once had no life, but then did. How? No one has a definite answer. Why? Definitely no one has an answer.
- As COVID-19 has taken hold of our planet, we've been hearing more and more reports about disordered sleep.<sup>456</sup> Scientists have found links between it and melatonin.<sup>7</sup> But though one of the symptoms of long-term COVID-19 is insomnia, no one knows for sure why that is the case, or what role exactly sleep plays in warding off or recovering from the disease.<sup>8</sup>

- In the 16th century, a Spanish nun, deep in prayer, sees an angel close by, small of stature, face burning, a beautiful cherub. In his hand, the cherub is holding a long golden spear, the point of which burns, too. He thrusts the spear into her heart, pulls it out and pulls out the entrails with it. It is terribly painful, but there's a sweetness in the pain. She moans. She is on fire with the love of God.<sup>9</sup>
- An old man is about to leave his little apartment when he realises that his keys are gone. He definitely remembers putting them on the hallway table and there's no one besides him living in the apartment. But after half an hour of searching, he still hasn't found them.

The distinction I want to make is this. Sometimes we mean by *mystery* a situation where a thing is not known that could or should be known. This is the case for three of the examples above: the circumstances of the young monk's murder, the link between COVID-19 and sleep and the whereabouts of the old man's keys are all things that we should, or at least could in theory, know. But we don't. And the fact that we don't but should or would ordinarily know focuses our attention on these unanswered questions. That's what makes them mysterious. I call this **the mystery of the unknown**. It is the domain of secrecy and mystery fiction.

But there is another kind of mystery in which the sought-after information is not only unknown, but in a sense cannot be known, or cannot be understood with the human intellect, or cannot be expressed in human language. This is what Teresa of Ávila experienced in her vision of the cherub and perhaps what you yourself experience when you think about the meaning and purpose of life. I call this **the mystery of the unknowable**. It is the domain of poetry, philosophy and religion.

What these two kinds have in common is of course that they have to do with knowledge, or more precisely its absence, together with a focus on or calling to attention of it. This latter part seems pretty important, because new information can be mysterious if it directs our attention to other things of which we have no knowledge. Mystery isn't so much the unknown or the unknowable as *pondering* the unknown or the unknowable. Odysseus Elytis, the great Greek modernist poet, has asserted, quite beautifully (though I don't know if he's right), that these modes vary from culture to culture:

There is a passage in my collection of essays where I say that Europeans and westerners always find mystery in darkness, whereas we Greeks find it in light, which for us is something absolute. To illustrate this I relate three memories. I tell of how I once, in the middle of the day, saw a lizard climb up on a rock (it wasn't afraid because I stood motionless, without breathing even) and then, in full daylight, with a multitude of small movements perform a veritable dance in honour of the light. On a separate occasion I experienced this mystery at sea, between the islands Naxos and Paros. Suddenly, I saw at a distance dolphins approach and pass us, leaping high over the surface of the water, at the level of our deck. The final memory is that of a young woman on whose naked breast a butterfly sat one day around dinner time as the cicadas filled the air with their loud creaking.<sup>10</sup>

It seems to me that the mysteries Elytis has in mind are those of the unknowable, the numinous, the transcendental. He is, after all, a poet, and poetry has in common with religious mysticism that it seeks to explain the sorts of things that cannot be expressed in words, only alluded to. Science is different. The sciences have comparatively little mysticism in them, because it's in the nature of the scientific method to ask the sorts of questions that can be answered and that can be comprehended intellectually. But in religious mysticism, knowledge of a divine sort is attained through means outside the intellect and outside human faculties of reason, through spiritual or intuitive means. "To have faith is precisely to lose one's mind so as to win God."<sup>11</sup>

## Mysteries of Cthulhu

*(There are spoilers ahead. If you have not yet read the story, this is a good time to do so. It's short enough to read in one or two sittings. You can find it [here](#).)*

*The Call of Cthulhu* tells the tale of a man – the narrator – who, finding “[a] queer clay bas-relief and [...] disjointed jottings, ramblings, and cuttings” among the papers of his grand-uncle, a professor of Semitic languages who has recently died under suspicious circumstances, follows in the deceased one's footsteps as he pieces together clues gathered from the dream reports of a young, unstable artist, press cuttings describing strange incidents from all corners of the world and a night raid in the swamps outside New Orleans led by an unwitting police inspector. These clues hint at obscure cults worshipping ancient beings known as the Great Old Ones. In the third and final act, the narrator comes across by chance a news report of the misfortunes of the crew of a New Zealand steam yacht, out of which only one man survived, an incident that leads the narrator to sense at last the full extent of everything that has happened.

Now you will see why I took that long detour before. I think that this story, in my view the finest that Lovecraft ever wrote, is extraordinary because it revolves around both kinds of mystery. There is the unknown, which is summoned by the narrator's gradually learning what his grand-uncle learned and then finding out what happened to the New Zealand steamer. There is the unknowable, which is the nature and history of Cthulhu, which nature and history, in the narrator's view and in the view of the implied author<sup>12</sup>, are the stuff of nightmarish horror, as is made clear in the story's famous opening paragraph:

The most merciful thing in the world, I think, is the inability of the human mind to correlate all its contents. We live on a placid island of ignorance in the midst of black seas of infinity, and it was not meant that we should voyage far. The sciences, each straining in its own direction, have hitherto harmed us little; but some day the piecing together of dissociated knowledge will open up such terrifying vistas of reality, and of our frightful position therein, that we shall either go mad from the revelation or flee from the deadly light into the peace and safety of a new dark age.<sup>13</sup>

This is the sort of knowledge that human brains are not fit to store or process. Now, Lovecraft was an atheist and a materialist.<sup>14</sup> But the implied author of *The Call of Cthulhu* has a lot in common with religious mystics. To them, the mysteries central to their religions couldn't be

transmitted by speech, but had to be intuited or felt. Lovecraft's gods likewise cannot speak with humans:

When, after infinities of chaos, the first men came, the Great Old Ones spoke to the sensitive among them by moulding their dreams; for only thus could Their language reach the fleshly mind of mammals.

Lovecraft evokes mystery in a variety of ways, the most common of which is also the least artful, namely that of simply *telling* us that something is mysterious, obscure, incomprehensible or indescribable. He does that a lot. But I don't think it's a flaw necessarily: I think these words are the foot soldiers in his war against the mundane. They are a necessary, if insufficient, tool, tiny particles of a shadowy cloud.

But he does have many more artful ways of doing it, too. One is through **anomaly** and **ambiguity**, in other words the quality of being either a member of a very rare category, or not an obvious member of any category. These qualities work to heighten the mystery because they are open questions, they point out something unknown. For instance in the police inspector's encounter with the swampland cultists:

There are vocal qualities peculiar to men, and vocal qualities peculiar to beasts; and it is terrible to hear the one when the source should yield the other.<sup>15</sup>

– the same cult-members which the narrator (and it may be worth pointing out here that Lovecraft had deep-seated prejudices) describes as “mongrel” and of “mixed blood”, highlighting their ambiguous nature. Of which origin are the cultists? Why are these people making sounds like animals? The young artist claims that R'lyeh, the sunken, “nightmare corpse-city” in which Cthulhu rests, in his dream-visions, had a geometry that was “all wrong”. This is later confirmed by the surviving seaman, who reports that the same city had a geometry that was “abnormal, non-Euclidean, and loathsomely redolent of spheres and dimensions apart from ours”. And so on.

Another technique that Lovecraft uses is **paradox**. A paradox is a seeming contradiction that may, upon investigation, not be a contradiction after all. A paradox poses a question, too: it asks, “If both parts of me are not true, which is wrong?” The Gnostics realised this and used paradox freely, for example the author of *Thunder, Perfect Mind*:

For I am the first and the last. I am the honored and the scorned. I am the whore and the holy. I am the wife and the virgin. I am the mother and the daughter. I am the limbs of my mother.<sup>16</sup>

Lovecraft does this, too. The Louisianan cult-members chant of Cthulhu's being both dead and dreaming at the same time. The tormented artist dreams of “a voice that was not a voice”. One of the cultists that were captured by the police inspector says of the Great Old Ones:

They had shape [...] but that shape was not made of matter. When the stars were right, They could plunge from world to world through the sky; but when the stars

were wrong, They could not live. But although They no longer lived, They would never really die.

All these things add up to a full-frontal assault on certainty and knowing. Lovecraft was a cynic; his cynicism was founded on the futility of being, on the insignificance of humanity in a vast cosmos.<sup>17</sup> Like the mystics, he believed that it was possible to pull down the curtain of appearances and reveal the true nature of things; but whereas they thought you'd see the deeper meaning of things, he thought you'd see life's meaninglessness. He was not the first person to suggest that this transcendental knowledge was not a bed of roses. Aldous Huxley wrote that "[t]he literature of religious experience abounds in references to the pains and terrors overwhelming those who have come, too suddenly, face to face with some manifestation of the *mysterium tremendum*."<sup>18</sup> Teresa of Ávila's ecstasy had equal measures pleasure and pain. But I think the mystics believed that this knowledge was ultimately good – it was, after all, knowledge of God.

Early on in *The Call of Cthulhu*, having recounted his grand-uncle's having correlated a number of dream reports, the narrator repeatedly notes that it's well that the dreamers lacked the opportunity to compare notes, or to hear of what the narrator himself has learned. Because this is knowledge that warps and shatters the human mind. He might have rejoiced that the ordeal was over. But as he prepares to recount the misfortunes of the surviving seaman, and having learned all there is to learn, the narrator issues a prophecy:

I shall never sleep calmly again when I think of the horrors that lurk ceaselessly behind life in time and in space, and of those unhallowed blasphemies from elder stars which dream beneath the seas.<sup>19</sup>

- 
1. Lovecraft, H. & Schweitzer, D. (2008). *Supernatural horror in literature & other literary essays*. Rockville, Md: Wildside Press. ↵
  2. My source here is [Wiktionary](#), though the Oxford English Dictionary and the Catholic Encyclopedia give similar origins. ↵
  3. This is the beginning of the plot of Umberto Eco's magnificent novel *The Name of the Rose*. ↵
  4. Li, Y., Qin, Q., Sun, Q., Sanford, L. D., Vgontzas, A. N., & Tang, X. (2020). *Insomnia and psychological reactions during the COVID-19 outbreak in China*. *Journal of Clinical Sleep Medicine*, 16(8), 1417–1418. ↵
  5. Huang, Y., & Zhao, N. (2020). *Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey*. *Psychiatry Research*, 288, 112954. ↵
  6. Casagrande, M., Favieri, F., Tambelli, R., & Forte, G. (2020). *The enemy who sealed the world: effects quarantine due to the COVID-19 on sleep quality, anxiety, and psychological distress in the Italian population*. *Sleep Medicine*, 75, 12–20. ↵
  7. Hamblin, [The Mysterious Link Between COVID-19 and Sleep](#). ↵

8. *ibid.* ↵
9. This scene is from the autobiography of Teresa of Ávila. ↵
10. From an interview by Ivar Ivask, translated into Swedish by Lasse Söderberg and back into English by myself. ↵
11. Kierkegaard, S. (1989). *The sickness unto death : a Christian psychological exposition for edification and awakening*. London, England New York, N.Y., USA: Penguin Books. ↵
12. I follow Wayne C. Booth in making the common-sense distinction between [the implied author](#) and the breathing (or, as it were, decaying) author. ↵
13. Lovecraft, H. & Joshi, S. (2017). *The call of Cthulhu and other weird stories*. London: Folio Society. ↵
14. Zeller, B. E. (2019). *Altar Call of Cthulhu: Religion and Millennialism in H.P. Lovecraft's Cthulhu Mythos*. *Religions*, 11(1), 18. ↵
15. Lovecraft, H. & Joshi, S. (2017). *The call of Cthulhu and other weird stories*. London: Folio Society. ↵
16. Meyer, M. & Funk. (2007). *The Nag Hammadi scriptures*. New York: HarperOne. ↵
17. Zeller, B. E. (2019). *Altar Call of Cthulhu: Religion and Millennialism in H.P. Lovecraft's Cthulhu Mythos*. *Religions*, 11(1), 18. ↵
18. Huxley, A. (2004). *The doors of perception and Heaven and hell*. New York: Perennial Classics. ↵
19. Lovecraft, H. & Joshi, S. (2017). *The call of Cthulhu and other weird stories*. London: Folio Society. ↵

# The Case Against Reality by Don Hoffman

Don Hoffman's book "[The Case Against Reality](#)" makes the argument that *reality as you know it does not exist*.

For obvious reasons, this is a difficult case to make. Hoffman does a better job than I could've expected.

But throughout the book, my mind was racing with doubts. In between the most convincing details of his argument, he's shunted entire chapters on barely-related research he's done, and name-drops his conversations with Francis Crick so frequently it's tough not to believe Don (or Don's editor) wants you to remember *Don* as much as he wants you to remember that reality is a sham.

Here I want to distill Don's argument, throw in a few thoughts of my own, and generally try to steelman the case against reality enough so that, by the end, it feels *obvious* that nothing you perceive actually exists.

In the second part, I'll ramble about what I found unconvincing in his arguments. And then I'll talk about why this book seems so important and relevant, *even* if huge chunks of it turns out to be false.

## I.

Consider, if you will, *color*.

If you think about it, color is really weird.

I remember wondering as a kid what color *atoms* were. This was before Google, so I ultimately just wondered and wondered until one day I saw it in a science textbook: a fuzzy, gray blur on a slightly darker gray background.

As was obvious in that moment, this image of an atom was not exactly a *photograph*. A photograph implies you have *so many* photons bouncing off of something and hitting the cells on the insides of the camera's sensor that you can actually *see* a consistent image. But for atoms, that can be a pretty tall order. Photons can get absorbed by an atom's electrons, or knock the electron out of orbit, or fly out at random times. And when I realized this – years later – I thought, "Wow! So atoms are smaller than color!" This sounded profound to me, since color obviously isn't a certain *size*.

But that's a bit of a figure-ground inversion. "Atoms" are not a special edge case where color breaks down. "Color" *is* a special edge case in which something reliably emits light 400-800 nanometers in wavelength at a human eye. A table or chair falls into the has-color category. A



single atom does not. Of *all* the information reaching our bodies at any given point, it is only that which is coded in a narrow band of electromagnetic radiation that we happen to perceive as hues.

Now I'm getting ahead of myself. I said *humans* can perceive light in the 400-800 nm spectrum, but other animals can perceive light outside of that band. To speak a bit loosely, they can see colors we cannot see. The birds and the bees can see UV (I know there's a pun in there, but I'm unable to find it). Certain cold-blooded animals can see infrared light. The electromagnetic spectrum technically has no upper or lower bound, so we essentially see a small sliver of a reality in which infinitely broad swaths are hidden from our eyes.

And, to be clear, it's not that we actually "see reality". It's more that sight in an *interface* with a small sliver of reality. But the better part of reality bounces along in what we could call darkness, since we have no interface allowing us to grasp any of it.

So, pop quiz. What's wrong with this statement? "Reality itself, if you could perceive it directly, is colorless".

Answer: "Perceive it directly" is a contradiction in terms. We perceive *everything* through our senses. That's what *perceive* means! We don't have direct access to *any* property of reality. Only what our senses have given us. Saying "reality is colorless" implies that perhaps it's grayscale, like that grainy image of an atom I saw in high school. But we're not talking about grays here – we're talking about a realm in which the very idea color is so misguided as to be a distraction. Asking what color any object "really is" is like asking what a mathematical equation really tastes like. [Mu](#).

And this applies to all of our senses. Reality isn't just silent – it's something in which the very notion of silence is a non-sequitur. To me, the fact that reality is not-even-silent feels a bit easier to grok than reality being not-even-colorless. And the other senses are even easier to recognize as *pure interface*. I don't know that anyone, when presented with the beatific vision under the hood of the universe expected it to *smell* like anything.

Realizing that senses are interfaces to tiny slivers of reality is a massive figure-ground inversion. They seem so fundamental. "There's no sound in space" feels wild the first time you hear it, but from our new vantage point, it should be obvious. There's no sound basically *anywhere* except the tiny planet you and your extended family evolved on.

Which is a nice lead-in. *Why* do we sense some colors, but not others? Why do we *hear*?

Fortunately, basically every "why" question in biology has the same answer: *selection*. Creatures perceive what is useful to their survival and reproduction. Those that didn't get selected by nature. And nature, it turns out, is a teacher who you *really* don't want to call on you.

So for us humans, this means seeing light in the 400-800nm range let our ancestors (a) live to reproductive age and (b) have more babies. Seeing *different* wavelengths or *more* wavelengths

didn't have that same advantage – it was energy and complexity wasted on something that didn't help with tasks A or B listed above.

(This doesn't mean we have zero interaction with different wavelengths of light. UV light will sunburn us – and eventually kill us of skin cancer. But skin cancer strikes *after* we have kids, which means our ancestors never experienced much selection pressure to see UV – and therefore we also are blind to it. Were we able to see it, I like to imagine we'd perceive strong UV with a creepy, eerie-looking glow, in the same way that other health hazards are likewise noxious – e.g. the smell of rotting meat, or the appearance of molding fruit)

And all the crazy animal senses that we can't comprehend – like sharks sensing electric fields and octopi sensing the polarization of light – those are all the same. A sliver of the properties of reality, morphed into something wildly different: *a conscious sensation that gets us to act in evolutionarily fit ways.*

Once you admit that we don't sense reality, but instead, our senses create an interface to the useful slivers of reality we need to survive and make babies, a lot of physics starts to make more sense. Or rather, the confusing things feel more expected. For instance, neutrinos are particles that so rarely interact with other matter, it took us 70 years from when they were theorized to find all 3 types in experiment. At first it feels a bit weird to think there are things in the universe that are so astoundingly shy. Why? If they don't interact with any known matter, what role can they possibly play? As I.I. Rabi said of the unexpected discovery of the muon, “who ordered *that?*”

But this line of thinking is like saying that once we've picked the low-hanging fruit (the part of the universe we can easily interact with), the top of the tree must be near. Again, invert figure and ground: the height of the lowest fruit is more a measure of *us* than the tree.

Why would we *expect* that everything in the universe readily interacts with us? What's not to say the largest fraction of reality remains undetectable and undiscovered, simply because it was never of any survival value for us to know about?

(I actually wrote the above sentence before thinking of dark matter and dark energy, but it's a perfect illustration of this idea. In the 1990s, we realized that literally 70% of the mass-energy of the universe was due to something outside of the realm of what we could see or interact with directly. Anyone would say that's hugely surprising, but if we've really intuited the idea that our senses are interfaces into a small sliver of reality, then it's no surprise at all. The height of the lowest-hanging fruit is not a measure of the tree)

Up to this point, we've only talked about *sensory* experiences. But this “interface to reality” principle applies to *all* our experiences. For instance, our experience of time, and our experience of 3D space.

Like our experience of vision, our experience of time feels pretty solid. At an intuitive level, time appears to march forward at an objective, steady rate. But Einstein shot down “objective” and “steady” a century ago, and whatever time is, it's not those.

(Specifically – and bear with me if you’ve heard this – time could speed up or slow down depending on how fast you were moving, or how near you were to very massive objects. The center of the earth is literally 2.5 years younger than the crust – not because it was formed first, but because *time moves slower down there*. And for a photon, ever-moving at the universe’s speed limit? The best we can posit is that it *doesn’t actually experience time at all*. Wherever it zips to and from, it’s always all of those places)

And this is just the well-established stuff! Physics is still in the process of figuring out exactly what time is. We can’t say with certainty. But the important thing to realize is that our sense of time (like any of our senses) is an *interface* with reality, and we were never promised our interface would capture the whole of underlying reality. In fact, the *purpose* of an interface is to *not* capture the whole. It’s to shunt away the extraneous details, and only communicate what’s *useful*.

In the absence of being able to describe what time actually is, I can only provide some guesses as to how different it might be from our experience of it. My favorite such guess comes from David Deutsch in his book [The Fabric of Reality](#), who compares time to a puzzle. A puzzle is meant to be viewed as a 2-dimensional rectangular image. But, we could also look at a puzzle in a different way. For instance, we could go from top to bottom, removing one row of the puzzle at a time (the first row left-to-right, the second right-to-left, the third left-to-right again, etc), and gradually concatenate a *single long row* with all puzzle pieces. Each piece would fit into the ones next to it, you’d have a completely deterministic series of pieces, you could always reliably reassemble the whole, and yet it would be an entirely silly way of looking at the puzzle.

Likewise, we experience time as a single long series of snapshots – but perhaps the most explanatory way to assemble those snapshots is some other structure. Despite being arranged in a different structure though, each individual snapshot *feels like* one in a long series. Why? Because even in a single snapshot, our brains contain memories of snapshots that affected this one in a way such that we call it the past. The past can exist concurrently, or all snapshots can exist in not-even-time, but from the vantage point of any one snapshot, we feel like there was a past and we are now beyond it.

Again, this is only somewhat idle speculation as to how things might actually work. For right now, we need only recognize that (a) our interface with time, like any sensory interface, will hide vast swaths of what’s actually there, showing us only what’s useful for natural and sexual selection, and (b) since we already know time is far more complex that it *feels* like it is, we should maintain wariness of any of our intuitions about it.

The next “sense” on the chopping block is our sense of 3D space – that of living on a fabric on which everything else plays out. Einstein definitively hacked apart time a century ago, but the cracks are only more recently starting to show in 3D space.

Theoretical physicist Nima Arkhani-Hamed says of it:

“Almost all of us believe that spacetime doesn’t exist, that spacetime is doomed, and has to be replaced by some more primitive building blocks”

But while no one yet knows what those primitive building blocks are, we should take the same lesson that we learned from our other senses: what we experience of 3D space is merely an evolutionarily useful representation that's misguidingly simple compared to the underlying reality.

A fun hunch as to what *is* more fundamental than space is from physicist Sean Carroll's book [Something Deeply Hidden](#). He notices that, of all the properties a particle might have, its *position in space* is the only one that allows it to interact with other particles. If two particles have the exact same velocity, cool – but indicative of nothing. If two particles have the same spin, wonderful – but again, it means nothing. If two particles have the same (or nearly the same) *position*, however – then they can magically influence each other's *other* properties. "Bounce off each other", in common parlance, but Sean would warn us not to read too much into the image of billiard balls ricocheting through space.

As Sean points out, there's one particular elephant in the physics room that messes with this idea, and it's *quantum entanglement*: the phenomenon by which two particles appear to influence each other's properties instantly and regardless of distance. To see this clearly, Sean recommends inverting figure and ground. Perhaps entanglement is *more fundamental* than space. Since nearby particles are always entangled, perhaps distance is just our way of *perceiving* entanglement – and, in the rare cases where entangled particles are incredibly far away, it breaks our notion of distance. But of course it does – because *distance* is merely the emergent mirage of a network of entangled particles.

And if "a network of entangled particles" sounds hard to grok, perhaps that's a sign it's only closer to the underlying truth. We spend all our time using an interface to reality. *Of course* that interface feels intuitive. Reality itself, by this measure, should almost surely feel strange.

In the end, reality is likely to have about as much in common with what we see, hear, feel, and experience as your computer desktop has in common with the billions of wires and transistors inside it. The purpose of the interface is not to *show* the underlying reality, but to *distort it usefully*. What we perceive is more like a trash can icon on your desktop. Of course there is no trash can; it's merely a useful interface for interacting with the billions of voltage levels your computer tracks as bits. The trash icon we see? Useful for sure, but not even remotely close to what lies underneath.

Put a spoon on your counter. When you look at it, it exists in your interface with reality – your sense of sight. But now look away. Is there still a spoon? If it has exited your interface, then only the underlying reality exists. A reality that we know has no color, no texture, no sense of weight, and doesn't live in time or 3D space. Whatever exists there is best not called a spoon. We only use that abstraction for the appearance of the spoon on our interface.

So this all begs the question: what exactly then *is* real?

Great question. Um. Let's change topics.

## II.

I want to talk for a minute about the parts of Hoffman's book I found unconvincing. Try as I might, there were a few things I had to omit from my steelman summary above, as I couldn't do them justice.

"When I have an experience that I describe as a red tomato a meter away, the content of that experience is not that there is – in objective reality, even when no one looks – a red tomato a meter away."

Hoffman is fond of talking about the disappearing tomato – one that ceases to be a tomato when no one is sensing it. Now I can get behind the "redness" of the tomato being interface, not reality. Indeed, that was my opener. I can even imagine the "meter" being a little shortcut of our perception, not a low-level concept of reality. But dang it, it's hard to imagine the tomato losing its "tomato-ness" just because no one's looking.

Frankly, the only analogy I can think of that makes sense here is software. Does a *simulation* of a tomato cease to be a tomato when no one looks at it? Well, yes. There still might be some file tucked away with the shape and texture of the tomato, but that file, in the end, is ones and zeros, likely implemented as higher and lower voltages on a sea of wires and transistors. And perhaps the operating system has seen to split up that file to be stored in chunks, not even localized in one part of the disk. At some point, we're talking about something enough layers of abstraction away from a tomato that the tomato-ness is just utterly gone.

And I suppose that's Don's implication? Perhaps we are akin to software beings (his video game example of choice is Grand Theft Auto, which I hope sets a precedent for all writing on the illusory nature of reality) realizing their experiences are merely an interface to an underlying reality, and even if we grok that "it's all made of zeros and ones!" we're still not *really* understanding what substrate we're on. We'd still be pretty surprised when, I don't know, a processor overheats and frames start dropping. You need to go deeper to understand that; even knowing "it's all ones and zeros" won't explain it.

Hoffman only mentions the idea that we're in a simulation a couple times in passing, and I'd pay to hear more of his thoughts on this. He is careful to avoid giving specific implementations of reality (entirely appropriately), but the lack of more in-depth metaphors for his thoughts can make it tough to understand exactly what he's implying.

"[A]toms and subatomic particles are not visible without special equipment, but they're still in space and time, and so they are still in the interface"

Or, to use his analogy:

"[This] claim is more like saying 'I know that the icons on my desktop are not the true reality. But if I pull out my trusty magnifying glass and look really closely at the desktop, I see tiny pixels. And those tiny pixels, not the big icons, are the true

nature of reality.' Well, not really. Those pixels are still on the desktop, still in the interface. They may not be visible without a magnifying glass, but they're part of the interface nonetheless."

Hoffman argues evolution shaped our perceptions – that is, our interface to reality – to be *useful*, not *truthful*. And that's a fantastic insight. If we expect that our perceptions (and therefore, our intuitions) about things are there to keep us *alive*, it causes us a lot less grief when those intuitions go haywire at, say, the quantum level.

However, the converse seems to be that evolution *wouldn't* shape our perception of things *not* useful for our own evolutionary fitness. Ignore atoms and electrons, which are undeniably critical for fitness. Let's look at something I mention above, something that seems a little less debatable: *neutrinos*. They are mysteriously reticent particles, slow to interact with anything else we've ever discovered. Understandably, it took our species a really long time to find them. I'm not positive of the following statements, but I would put some money on (A) neutrinos were never at any point critical to human evolution, and (B) we have no way of sensing them in our evolutionary environment.

To me, this is an indicator that they're *not* part of our interface with reality – they're actually a step closer to what *underlies* our interface.

But if we can sense them *at all, even with our fancy tools*, Hoffman says they're part of the interface, not reality. I'm not sure how he arrived here.

Another example: the spin of an electron. I like this in particular because it's so inconceivable to our intuition. In college, we were simply told "Yeah, electrons have a property called 'spin', it's measured in units of angular momentum, it can be positive constant, negative constant, and nothing else, and by the way – *it's useless to try and imagine the electron as actually spinning*". Again, I don't believe electron spin was ever fitness-enhancing to perceive. But we've built tools that *can* perceive it. And we've built math that can predict it, even if we can't explain it per se. That some weird property from beyond the pale leaks into our environment seems like a knock on the door from the next level down, not simply "pixels viewed through a magnifying glass". Hoffman hints as much when he talks about the holographic principle – the idea that the maximum amount of quantum information (never mind precisely what this is) storable in a volume is proportional to the *surface area* of the volume, not the... you know... *volume* of the volume. So what inconceivably weird aspects of our environment are hints to reality, and which are pure interface?

Until I hear otherwise, I like the converse of Hoffman's own metric: if it wasn't fitness-enhancing to sense, it's not part of the interface.

"The network dynamic of conscious agents is compressed into dynamics within spacetime. For instance, perhaps a dynamical evolution of conscious agents toward small-world networks may appear in spacetime as the dynamics of gravity"

So ends the very last paragraph of the appendix, and Hoffman transitions directly into thanking such luminaries as Daniel Dennett, David Chalmers, Steven Pinker, V. S. Ramachandran, and... Deepak Chopra? I guess enough anti-reality talk gets you mixed up with an interesting crowd. But no hate. I'll explain in a minute, but *someone's* got to be thinking these thoughts.

Anyhow, the last sentence of this book was like a needle going directly into a tire that we've been inflating for over 200 pages. Hoffman has walked us through the illusory nature of all our perceptions, including those of space and time, decried even the objective existence of *atoms*, and finally, while we're at this absolute trough of epistemological certainty, he sneaks in a few of his own suspicions about Conscious Realism – the idea that conscious agents are more fundamentally real than the supposedly objective world around us. And like I said above, fair enough. I want someone to be thinking through this stuff.

But then in the very last sentence, we get a glimpse into how he suspects conscious agents might work, and it's... gravity is a side effect of conscious agents needing to network with each other?

I so badly wanted to know what he actually thought was going on down there in the basements of reality. And now I suddenly don't.

But.

### III.

I remember reading once a plea by a physicist: *we really need to figure out what to teach the next generation of future physicists so that they can figure out the stumbers in quantum mechanics*. I believe it was said with an eye towards making video games that have locality-breaking or cancelling-probability elements. But Don Hoffman's philosophy, more than anything I've ever read, feels like the answer here.

What do we teach young scientists so they can see farther than us? Teach them their intuitions of reality *owe them nothing*.

A lot of questions in science now revolve around things that are very hard – or currently impossible – to interact with directly: dark matter, dark energy, the multiple worlds interpretation of quantum mechanics, consciousness, existence itself. When facing questions about a universe that we can only detect a small fraction of, it's tempting to sweep ideas away wholesale. "There's no hard problem of consciousness", or "MWI isn't even science" are not uncommon opinions. But Don's viewpoint is that *of course* the universe is filled with structures that are nigh-impossible to observe. Our senses were shaped by evolution to allow us to see what we *interact* with – so all the parts of the universe we *can't* interact with will obviously be dark. But nothing suggests that they don't exist; merely that we won't be able to see them directly.

And where we can't see and can't interact, we won't have a good intuition for what's there. So: best to expect our intuitions will be trampled roughshod.

I think the most important part of *The Case Against Reality* is this: *before reading it, certain results would feel weird because they weren't explained by current theories; after reading it, current theories feel weird because they can't explain certain results.* This book is not about consciousness, and it's not about quantum gravity, and, well, OK, it *is* about whether our senses have any pertinence to the fundamental nature of reality, but even if it weren't, it would've made me realize that the gaping holes in our explanations of our environment won't be filled by *normal-sounding theories*. That ship has sailed. Whether that's the hard problem of consciousness, or the unification of relativity and quantum mechanics, or what will replace spacetime, Hoffman has taken a look at those problems and realized our intuitions owe us no hints in truth-seeking. But why should they? That was never their job.

Here are a few things *The Case Against Reality* has changed my mind on – or, if not changed, then somewhat strengthened the direction I've been leaning:

Before: "The fact that anything exists at all feels crazy and almost unbelievable" After: "Oh shut up. You don't know the half of it, or perhaps the hundredth of it, or perhaps the  $10^{30}$ th of it! Maybe literally *everything* exists. The only thing you know about the stuff you can't see is that your intuitions aren't going to be much use!"

Before: "We can't see (or meaningfully interact with) other universes in the multiverse. The Multiple World Interpretation of quantum mechanics is really tough to believe!" After: "Hey moron, wasn't it evolution that gave you sight? If you can't *interact* with something, it's not affecting your fitness. You should have zero intuition about what's beyond the realm of your senses. Zero. What on earth does *not seeing something you can't interact with* count for its non-existence!? You're talking in circles."

Before: "Some particles move *backwards in time*!? What even *is* that!?" After: "You really don't get it, do you? Why do you think your perception of *time* has any more relation to reality than your perception of *color* does to photon wavelengths? Time is a useful interface plastered over whatever it is that we're living in, but don't expect it to be too intuitive beyond the realm in which it's useful to your evolved senses."

Before: "Atoms are smaller than color. How poetic and weird!" After: "You know is outside the realm of color? The literally 95% of the universe that we have no current means of detecting directly, and yet are sure exists. Color is the edge case, man. A tiny peep-hole into a little sliver of dark existence. Come on."

Before: "Consciousness is the hard problem! So hard! Everything else seems simple by comparison" After: "Dude, if you haven't figured out the hard problem, don't get too confident that you actually understand all the easy ones. Remember, from here on out, the going gets weird. We've picked all the low-hanging fruit from the tree of theories. We're sailing well past our intuition at this point, and what we find, should we find it, might re-write everything we thought we already understood."



The history of science feels like one slap in the face after another to anthropocentrism. We're not the center of the universe. Not even the center of the solar system, let alone the galaxy – of which there are, by the way, billions. And your great-great-grandparents were one-celled organisms. And there may be infinitely many copies of you. You thought you were a special snowflake, and you were right – at least in the sense that there are billions of you!

And now as we run up against the edges of time and space and matter and consciousness, we'll have to view our own perceptions – and the intuitions they foster – not as a reliable guide to what's really there, but as such a nearsighted peering through such a dim, narrow keyhole that the shadows we see on the other side are practically distractions to understanding what it is that actually lies there.

In any case, good luck to us! Seems like we'll need it.

[The Case Against Reality](#) by Don Hoffman

# The Cult of Smart: How Our Broken Education System Perpetuates Social Injustice by Fredrik deBoer

Education is a huge topic. Too huge, really, because almost everything we care about, as humans, has an element of inculcation--of learning. We are great imitators; it is the [secret of our success](#). Without education, we're little more than naked apes, so when you talk about education, you are in some sense talking about *the thing that makes us human beings*.

Classroom education (itself a subset of "formal" education) is a slightly more manageable topic, albeit in much the way that some infinities have lesser cardinality than the infinities containing them. In the United States, formal education arguably begins in 1635 with the "public" [Boston Latin School](#), though attendance was at the time neither free nor compulsory; [Harvard was founded the following year](#). In the 1640s [Massachusetts followed up](#) with several laws holding parents and communities responsible for the education of children (particularly in literacy), but these laws did not require *classroom* education and were not, as far as I have been able to determine, very strictly enforced. It was more than 200 years before [Massachusetts became the first American state to levy fines](#) against parents who did not send their children (aged 8-14) to a classroom most days. If you've studied education at all, there's a good chance you've heard names like [Horace Mann](#) and [Henry Barnard](#). These men witnessed, in the 19th century, a nation in turmoil (remember, the Civil War breaks out in 1861, after decades of increasingly acrimonious partisanship over questions of slavery). Their proposed solution was to create social harmony by inculcating social values in the rising generation, a mixture of literacy and numeracy with Christianity and "common public ideals."

A republican form of government, without intelligence in the people, must be, on a vast scale, what a mad-house, without superintendent or keepers, would be on a small one.

Over 150 years later, a lot has changed--and yet, perhaps not as much as sometimes seems. In her 1987 manifesto, *Democratic Education*, Amy Gutmann (now president of the University of Pennsylvania) wrote,

We disagree over the relative value of freedom and virtue, the nature of the good life, and the elements of moral character. But our desire to search for a more inclusive ground presupposes a common commitment that is, broadly speaking, political. We are committed to collectively re-creating the society that we share. Although we are not collectively committed to and particular set of educational aims, we are committed to arriving at an agreement on our educational aims (an agreement that could take the form of justifying a diverse set of educational aims and authorities). **The substance of this core commitment is conscious social reproduction.** As citizens, we aspire to a set of educational practices and authorities to which we, acting collectively as a society, have consciously agreed. It follows that **a society that supports conscious social**

**reproduction must educate all educable children** to be capable of participating in collectively shaping their society.

This is about as good a summary as one could hope to get of what is sometimes called "liberal education." Liberal education presupposes a mutual commitment to coexistence, and has future coexistence as its overriding aim. This is more complicated than it might seem; people who fail to achieve basic literacy are arguably locked out of our mutual project, people who seem to reap no benefit from the project may think they have little reason to support it, people who do benefit and participate might *overlook* the extent to which it is the project (rather than, say, their own intellect) that has given them the life they enjoy, etc. Peaceful coexistence is always a work-in-progress. This may be part of what led [Paul Goodman](#) to opine that

The compulsory system has become a universal trap, and it is no good. Very many of the youth, both poor and middle class, might be better off if the system did not exist, even if they had no formal schooling at all.

[Freddie deBoer agrees](#), more or less. Some reviews of *The Cult of Smart* argue that it is a less sophisticated rehash of Charles Murray's 2009 [Real Education](#) (yes, [that Charles Murray](#)), or point to an overlap between deBoer's concerns and the ones Byran Caplan made in 2018's [The Case Against Education](#). These are both plausible points of comparison, but in some ways simply too *new*; to understand the depth of the well from which deBoer is drawing, a greater sense of history seems required. The new vocabulary, research, and (perhaps especially) biological understanding from which Murray and Caplan draw do not lead them to conclusions all that different from Goodman's, just as a century-plus of educational reforms did not lead Gutmann to dramatically different conclusions as those drawn by Barnard and Mann. So how does deBoer fit into this mess, and what does he bring to the crowded table? At the risk of spoiling the rest of my review, the answer appears to simply be "communism."

The introduction of *Cult* is vaguely autobiographical. DeBoer vignettes some negative experiences he and others have had with American education, and then he alludes to the possibility that this is a function of heredity: some people are better biologically-equipped to succeed in school than others. He directly quotes Scott Alexander's [Parable of the Talents](#) in explaining that recognizing differences in talents is entirely compatible with a "belief that all people deserve material security and comfort." DeBoer's complaint is that schools are sorting mechanisms used to parcel out success in an intellectual meritocracy, and that this excludes some people from living the good life. Or maybe his complaint is slightly different, something like "education was supposed to reduce inequality, but it doesn't."

There are interesting moral arguments that one is equally culpable whether one *causes* a harm, or fails to cure it, so if this is a mistake, it is at least not a mistake unique to deBoer. But at a purely practical level, "schools cause inequality" is a *very different claim* than "schools fail to fix inequality" because each complaint implies very different solutions. If public education *causes* objectionable inequality, for example, then simply *abolishing public education* would be a plausible response. But if schools fail to *fix* objectionably inequality, then "so what, that's not something schools are capable of fixing" might be a plausible response. That these are really two

very different complaints is not something deBoer particularly addresses; he seems content to identify *any* plausible complaints against the liberal status quo.

As an aside, at the risk of sounding *incredibly* snobbish, I have to say: the fact that deBoer purports to attack liberal education as an egalitarian pursuit, without so much as *mentioning* Amy Gutmann, raises serious doubts about his merits as a scholar. He addresses Locke and Rawls (even if a bit shallowly), so I wouldn't necessarily assign him a *failing* grade on the matter--but Gutmann is the [highest paid university president in the Ivy League](#), and her contributions to the idea of egalitarian liberal education are in no way niche or obscure.

But the point may be moot; even *had* he cited to Gutmann, the outline of deBoer's argument would probably not have changed. Through the first seven chapters, about 2/3rds of the text, it looks something like this:

- The ability to succeed in school has become a primary distinction between haves and have-nots.
- Public education purports to reduce inequality, but as education has become more ubiquitous, inequality has actually increased.
- Public education does not create "equality of opportunity" because it cannot address inborn inequalities.
- "School quality" is not especially relevant to anything; it neither improves equality nor even especially improves individuals.
- Differences between individuals are predominantly inborn.

Suppose you accept all five points: can you derive any necessary conclusion from them? I certainly can't. Some of these points have been made more thoroughly, or more persuasively, by folks like Murray and Caplan, and more broadly they seem to be a contemporary re-tread of Goodman. I think each point has merit. But what deBoer seems to expect is that, once we've accepted all these points, we will see that "liberal education" is a failure. Our goals ("equality" is the ill-defined goal deBoer seems to assume his readers share with him) cannot be served by the status quo, and so we will be ready to

truly reconcile our egalitarian impulses with the reality of genetic predisposition, . . . to remake society from top to bottom, in schools especially but throughout our systems from birth to death.

This simply *does not follow*. Perhaps our "egalitarian impulses" extend only to equal treatment under the law, or to equal dignity and respect, or to equal access to public goods, or any of a thousand other egalitarianisms that do not rise to the level of preferring *equality of outcomes*, as deBoer explicitly does. His criticism of American public education seems basically cogent, if occasionally incomplete or, perhaps, symptomatic of motivated reasoning. But when he observes that

We sink vast sums of money into quixotic efforts to make all of our students equal

it does not seem to occur to him, at all, that we could therefore choose to *stop doing that*. Instead, bizarrely, he recommends we *continue* doing that--indeed, he thinks we should *pay teachers even more money* to keep doing that. Only instead of trying to make students equal by teaching them *math*, we should make them equal by teaching them to care about one another, to be compassionate, to work to the best of their abilities and be grateful to receive from others in accordance with their needs. Why deBoer thinks schools will be any better at teaching children these things, than they are at teaching children math, is never expressed or explored. Why deBoer fails to notice that there is no reason, in principle, to think that people's *dispositions* are any less governed by their DNA than are their *capabilities*, I can only guess, but it is an absolutely *glaring* oversight. What do we do, in his perfect world, with children who are predisposed to be *bad at caring*? What do we do with teachers who are bad at teaching it? DeBoer seems to be laboring under the delusion that teaching people to *behave* is substantially less quixotic than teaching them *algebra*.

Well, having described the problem as he sees it, deBoer devotes the final two chapters of the text to solutions. One chapter is a list of "limited reforms that would still do a great deal of good for students and teachers." Of these, one (universal childcare) has no obvious connection to public education, unless deBoer is trying to say that public educators are really just babysitters who should be treated as such. One is a cherry-picked whinge about charter schools (which deBoer seems more likely to detest because they are a form of private property than because there is anything uniquely objectionable about them). And three (lower the dropout age, loosen standards, and stop emphasizing college) are variations on a theme: "increase equality by lowering your expectations." I am skeptical of the benefits of universal childcare but not strongly *opposed*; I simply don't see its relevance to deBoer's project. Likewise his rant against charter schools is obviously not *unrelated*, but still struck me as a significant red herring. Rather, his only truly topical proposal--lower expectations--strikes me as exactly the *wrong* way to deal with children. I don't know how many children deBoer has raised to adulthood, but I've been through the process a couple of times and never seen anything to persuade me that *lowering* my expectations is a productive way to interact with them. But since deBoer himself seems to think that even these reforms cannot save us from "an Eloi and Morlock future where the college educated . . . pull further and further away," it is not obvious that there is anything further to be gained by meditating on this list.

In the final chapter of *Cult*, deBoer explains why communism is *just so great*.

The amount of second-hand embarrassment I felt while reading this chapter was excruciating. If you've ever read Upton Sinclair's *The Jungle* or Ayn Rand's *Atlas Shrugged* you may already have some idea what I'm talking about--in those novels, there comes a point where the author seizes the narrative to preach directly at the reader through their characters. It's graceless and uncomfortable even if you happen to agree with the message. *Cult* inverts the technique--deBoer's is a work of nonfiction that ends with a saccharine short story about how great life could be, if only we were all communists. A short, *fictional* story--why deBoer didn't share a *true* story from one of the many *actual* communist countries that have existed over the past hundred years, I leave as an exercise for the reader. Also in this chapter: effusive praise for Obamacare, advocacy for student loan forgiveness (even though it is "not a progressive expenditure"), and a

call for job guarantees and universal basic income. What does any of this have to do with our supposedly-broken education system?

It seems to me that the *Cult of Smart* is best understood as two unfinished texts, inartfully mashed together by an essayist with no serious experience crafting long-form arguments. In the first book, the shortcomings of public education in 21st century America are observed. To finish this book, one would need to consider the *strengths* of public education in 21st century America, and then weigh the costs of making particular alterations to the status quo. Can we do better with more spending? Can we do the same or better with *less*? This might be a primarily empirical inquiry, or a mostly theoretical one, but either way it would need deeper research and analysis than deBoer ever manages to summon. What would Amy Gutmann's *Democratic Education* or Caplan's *Case Against Education* look like, if they had been written by Marxists?

In the second book, education is just one consideration among many pointing toward communism as a solution to the harms brought about by human biodiversity. Once a person accepts that human biodiversity ensures that some lives are going to go better than others, one might conclude that this is good reason to order society in ways that alleviate the burdens of the worst-off. [Prioritarianism](#) is a form of (or arguably a supplement to) egalitarianism that fits approximately this description, and perhaps a case could be made that prioritarrians should favor political communism. Or maybe something straightforwardly Marxist would be more up deBoer's alley. It is harder for me to envision the contents of such a book, since I could never myself write it, but I assume that a chapter or two would need to be devoted to the primary role of schools as centers of political indoctrination rather than as centers of qualitative and quantitative inculcation. What does "cultural reproduction" look like to a communist who preaches anarcho-syndicalism? What would public education look like, if Mann and Barnard had been Russian Leninists instead of American Christians?

But deBoer wrote neither of these books. Instead we get a scattered mess. It is at most a list of grievances appended to a list of preferences, with scant connection drawn between them. DeBoer is a master essayist, but his magic appears to tap out around 2000 words. Which is too bad, really; it seems to me that the U.S. could use some thorough, intelligent education reform, and that's more likely to happen if progressives and conservatives can find some common ground on which to build compromise solutions. But if there is anything deBoer avoids more studiously than clarity, it is compromise.

In a sea of red  
five yellow stars shine brightly.  
This book gets just one.

# The Culture of Narcissism by Christopher Lasch

*The Culture of Narcissism* is a book that prophetically identifies a bunch of very twenty-first century cultural traits while also providing a bizarre and sometimes self-contradictory account of how they came into being. This is a largely because about half of it goes into long tangents about psychoanalytic theory citing Freud and Melanie Klein and so on, and accommodating psychoanalysis into anything usually comes at a cost to coherence. *The Culture of Narcissism* therefore consists of a set of mostly interesting complaints about contemporary society and a set of mostly wild explanations for them.

Despite being published in 1979, the complaints are particularly fitting for the impact of the online world. The problem is that people are getting more angry with each other, more stupid, more inclined towards immediate gratification, more self-regarding, and generally more useless at life. Lasch describes 'a type of personality that ought to be immediately recognisable, in a more subdued form, to observers of the contemporary cultural scene: facile at managing the impressions he gives to others; ravenous for admiration but contemptuous of those he manipulates into providing it; unappeasably hungry for emotional experiences with which to fill an inner void; terrified of ageing and death.'

Lasch offers two main explanations for this: one is that unfettered capitalism atomises society and erodes previous networks of social relationships, a void then filled by corporate and state bureaucracies which in turn makes people grasping, feckless and incompetent. The other is that in the absence of sound paternalistic authority figures, something goes developmentally wrong and we all end up with aberrant psyches, which in turn shape new, undesirable cultural norms which skew towards the short-term pursuit of self-interest at the expense of pretty much everything else. The first of these explanations is why Lasch is popular among [post-liberals](#), who are over the "double liberalism" of economic neoliberalism and social progressivism, both of which they see as corrosive to the bonds that maintain a favourable human environment, and propose bringing back Culture in an economically left-of-centre, neoreactionary-lite kind of way.

Lasch writes that 'Every society reproduces its culture – its norms, its underlying assumptions, its modes of organising experience – in the individual, in the form of personality. As Durkheim said, personality is the individual socialised. The process of socialisation, carried out by the family and secondarily by the school and other agencies of character formation, modifies human nature to conform to the prevailing social norms.'

Even if the evidence suggests that people get socialised more by their peers than their family, this sounds reasonable enough so far. But that's not how the argument continues. 'Each society tries to solve the universal crises of childhood – the trauma of separation from the mother, the fear of abandonment, the pain of competing with others for the mother's love – in its own way, and the manner in which it deals with these

psychic events produces a characteristic form of personality, a characteristic form of psychological deformation, by means of which the individual reconciles himself to instinctual deprivation and submits to the requirements of social existence. Freud's insistence on the continuity between psychic health and psychic sickness makes it possible to see neuroses and psychoses as in some sense the characteristic expression of a given culture.'

Lasch takes psychoanalytic theory very, very seriously. He is clear that narcissism means mass literal pathological narcissism in a strictly Freudian sense, not metaphorical narcissism in the sense of people being unduly selfish and self-involved, or a social malaise with vaguely narcissistic characteristics. According to Lasch, 'Theoretical precision about narcissism is important not only because the idea is so readily susceptible to moralistic inflation but because the practice of equating narcissism with everything selfish and disagreeable mitigates against historical specificity. Men have always been selfish, groups have always been ethnocentric; nothing is gained by giving these qualities a psychiatric label.' He then proceeds to give a psychiatric label to all contemporary Americans.

This is frustrating. It also undermines his more credible argument about the structural impact of hyperliberalism: if the one true mechanism for how people become narcissists is actually something about bad breasts and absent patriarchs, what does it matter if, say, schools are dysfunctional? And it adds long passages of boring and entirely unsubstantiated psychoanalytic digression that you have to skip through to stay sane between the complaints about the decline of Man.

\*\*\*

Complaint No. 1 might be summed up as 'everyone is a bit useless':

'Having surrendered most of his technical skills to the corporation, [the average American] can no longer provide for his material needs. As the family loses not only its productive functions but many of its reproductive functions as well, men and women no longer manage to raise their children without the help of certified experts. The atrophy of older traditions of self-help has eroded everyday competence, in one area after another, and has made the individual reliant on the state, the corporation, and other bureaucracies.'

This seems reasonable. It's the kind of thing the Unabomber might agree with, except he would argue that the correct response to the technologising push towards corporatism and away from individual autonomy would be to go and live in the woods and maybe bomb the odd software firm, while Lasch doesn't really have a response. It's trickier for Lasch because his anti-capitalist sympathies presumably entail more reliance on the state, not less, at least economically.

Complaint No. 2 might be summed up as 'having no boundaries is bad for social relations':



‘Conversation takes on the quality of confession. Class consciousness declines; people perceive their social position as a reflection of their own abilities and blame themselves for the injustices inflicted on them. Politics degenerates into a struggle not for social change but for self-realization.’

Most of this makes sense apart from the people blaming themselves for injustices and their social position. That definitely isn’t accurate now, even if it was previously. Class might be the last aspect of systemic injustice to be noticed, at least in American politics, but systemic injustice is definitely both noticed and the object of both social change and cynical self-realisation.

The idea of a culture of narcissism applies especially well to the online social world. One of Lasch’s bugbears is the erosion of boundaries in social relationships and the tendency towards emotional oversharing. This seems particularly true now: social media rewards emotive content with engagement and validation, and the phenomenon of the confessional form in everything from autofiction to memoir to journalism has exploded, especially among women writers. But is this a product of pathological narcissism or a product of a set of incentives that alter our behaviour? Writing about yourself world is easy, quick copy, and therefore cheap copy in publishing and journalistic spheres financially hollowed by technology: you can bash out a few thousand words much faster when you don’t have to research or plan it, and people find it relatable. You could maybe argue that people seeking relatable emotion-bearing content is ‘unappeasable emotion-seeking’, but I don’t think it’s that so much as low-hanging fruit, as easy to consume as it is to produce – it doesn’t entail any imaginative work to access rewarding feels.

A side-effect of having these incentives is that people who probably do qualify as pathologically narcissistic get more airtime from them than they might otherwise have done. But that isn’t an argument for everyone suddenly being pathologically narcissistic: it’s more that the people who are now have a way of being very, very visible. Most people oversharing on social media are just acting in ways that appear to offer immediate gain in terms of engagement and validation. Whether it does actually succeed in making people feel more valid and connected is questionable; emotion-heavy discourse, as Lasch points out, makes for hostile, rancorous relations.

Complaint no. 3 is basically ‘People are getting more stupid’:

‘Sweeping social changes, reflected in academic practice, thus underlie the deterioration of the school system and the consequent spread of stupidity. Mass education, which began as a promising attempt to democratize the higher culture of the privileged classes, has ended up by stupefying the privileged themselves. Modern society has achieved unprecedented rates of formal literacy but at the same time it has produced new forms of illiteracy. People increasingly find themselves unable to use language with ease and precision, to recall the basic facts about their country’s history,

to make logical deductions, to understand any but the most rudimentary written texts, or even to grasp their constitutional rights.'

Lasch is quite straightforward on what is wrong with education, which is that the watered down one-size-fits-all version of traditional education doesn't do anything well. He doesn't have any suggestions on improvements; basically everything is problematic either for being too elite or not elite enough. In particular, education is indulgent and demoralising, avoiding the potential confrontation of allowing students to experience challenge or failure and passing everyone regardless of effort or ability.

Complaint no. 4 is that parenting is too permissive:

'Feeding schedules gave way to feeding on demand; everything now had to be geared towards the child's "needs." Love came to be regarded not as a danger but as a positive duty. Improved methods of birth control, according to the progressive creed, had freed parents from the burden of raising children, but this freedom in practice seemed to boil down to the obligation to make children feel wanted at every moment of their lives.'

Lasch mentions 'debased versions of Freudian theory' here by which I assume he means attachment theory; it isn't clear what the correct Freudian theory says on childrearing. Parenting is also obsessed with authenticity: parents are encouraged to share their true feelings about any given situation instead of making a more detached, objective statement about it and children are free to express their negative emotions in the spirit of self-actualisation. This, Lasch says, 'reflects the collapse of parental guidance and provides it with a moral justification. It confirms, and clothes in the jargon of emotional liberation, the parent's helplessness to instruct the child in the ways of the world or to transmit ethical precepts' – instead, a panoply of outside experts take over in both state and commercial sectors.

This is only part of the picture, though. The main problem for Lasch is that parent and authority figures are devoid of authority – but he never details exactly what cultural code creates the correct kind of parental authority. If a family or social institution has an authoritative role regarding their charges, that must mean that they are doing some kind of behavioural shaping. Lasch never clarifies what this shape is supposed to look like; all we know is that it is no longer there. Perhaps this is because Lasch doesn't actually think cultural codes around parenting or authority are that important, or that if they are it's because they are the product of paternalistic authority, which has a character-forming magic in itself by means of some weird Freudian mechanics.

Overall, the issue isn't the lack of a cultural code or superego so much as that we now have the wrong, bad sort: 'The decline of institutionalised authority in an ostensibly permissive society does not, however, lead to a "decline of the superego" in individuals. It encourages instead the development of a harsh, punitive superego that derives most of its psychic energy, in the absence of authoritative social prohibitions, from the destructive, aggressive impulses within the id.' By this I assume Lasch foresaw people screeching at each other on Twitter, which is generally a good demonstration of

everything that is bad both in *The Culture of Narcissism* and in the world more generally.

\*\*\*

If there's a swathe of people online oversharing emotional content that probably won't serve them well long-term, regularly picking fights with other people and posting stuff that crudely signals certain status-signifiers, that sounds as much like a bunch of five-year-olds jostling for status in the playground as pathological narcissism.

And an alternative to saying, 'ah, that may be true but that's what you get from arresting the correct individuation of the ego' is to wonder if the reason people seem to act like increasingly like five-year-olds is that five-year-olds don't have much executive function. Similarly, if lots of people are inclined towards self-gratification, lack emotional boundaries and underperform intellectually because learning involves too much effort, maybe that's also more of an executive function problem than a literal pathological narcissism problem. Albert Ellis, the progenitor of Rational Emotive Behavioural Therapy, a type of CBT, might have described Lasch's various complaints as examples of low frustration tolerance – some examples from the Frustration Discomfort Scale via [Wikipedia](#):

1. Emotional intolerance, involving intolerance of emotional distress.
2. Entitlement, involving intolerance of unfairness and frustrated gratification.
3. Discomfort intolerance, involving intolerance of difficulties and hassles.
4. Achievement, involving intolerance of frustrated achievement goals.

These are the same problems that Lasch identifies; Ellis's [three core irrational beliefs](#) fit Lasch's diagnosis of narcissism perfectly, but Ellis doesn't even see them as pathological, just as a very human tendency to think unreasonably when circumstances allow. Ellis's solution to this would be that people need to be a bit more reasonable and get better at tolerating frustration, and that this can be learned in therapy. If Lasch is correct in observing a society-level tendency towards these behaviours, it's a similar complaint in many ways to [The Coddling of the American Mind](#) which sees them through a CBT lens. And if that is the case, maybe a postliberal-style society-level solution is a more workable idea than everyone being in therapy, even if it is a less batshit form of therapy than psychoanalysis.

The postliberals propose using the state to encourage and maintain a healthy ecology of tight, local social networks. [David Goodhart describes postliberalism](#) as seeing 'people as embedded in relationships, and wider groups, and conceives of their wellbeing as being dependent on those relationships and the state of the wider communities they are part of. One might call this embedded individualism. Freedom does not already exist inside each individual, it has to be created.'

This goes against the conventional liberal assumption that you get indefinitely freer as constraints lessen. But maybe they have a point – it's hard to act autonomously if you

lack the tools of self-control that allow you to operate beyond immediate gratifications and discomforts. One of the things that lockdown has shown us is that [people get worse at basic self-care like eating healthily and getting enough exercise](#) when we're cut off from real-life contact with other humans. I don't think this is because of insecure maternal attachment or being terrified of death looming round the corner dressed as Covid – it's that food is nice and comforting and exercise is hard and boring, and without the incentive to look okay for others, food is more likely to win out. If you ditch the constraints of social norms it's harder to put the effort in. Maybe this is why loneliness is toxic: we rely, at least in part, on other humans and their behavioural expectations to self-regulate successfully, and self-regulation isn't just good for social relations. It's good for our own wellbeing too.

If you do what Lasch quite explicitly thinks is a bad thing, i.e. treat the overall framework of narcissism, superego, id etc very loosely and metaphorically rather than as Freud's gospel truth, *The Culture of Narcissism* makes more sense. In a parallel-universe version of this book, what Lasch describes isn't a culture of narcissism but a culture of low frustration tolerance. The superego works as a set of internalised social constraints on individual id-type desires, like a social domain of executive function. Without a set of shared cultural norms around behaviour which people are socialised into, that social domain of executive function gets lost at a cost to our overall executive function, and we behave more childishly. The top-down mechanism of too much capitalism and too much social liberalism making people incompetent and dependent makes sense, and the bottom-up mechanism of permissive parenting and education making people emotionally labile and entitled makes sense.

There are points where Lasch touches on this: 'According to Henry and other observers of American culture, the collapse of parental authority reflects the collapse of "ancient impulse controls" and the shift "from a society in which Super Ego values (the values of self-restraint) were ascendant, to one in which more and more recognition was being given to the values of the id (the values of self-indulgence)."' But his attachment to Freudian dogma stops him from developing it in an interesting way.

Even if the authority enforcing superego values turns out to be more social than parental, the complaint about the collapse of ancient impulse controls is one we see across history, and especially at points where civilisations do well enough to relax their more tiresome constraints that people don't enjoy and edge onto the cusp of decline. That's when you get Socrates or Tacitus or whoever lamenting the intellectual and moral laxity of the youth and declining public standards.

Creating superego values lies not in facilitating maternal attachment and psychic individuation so much as in imposing constraints and demands and, in doing so, scaffolding frustration tolerance. A society that demands as its social norms the restraint, for example, of negative emotions in the service of civility, or the obedience of its population in the dull, grinding service of certain societal goals, is a society dependent on individual-level ability to maintain a degree of emotional continence and persistent effort. But maybe there are individual-level benefits to this too: if you

normalise high frustration tolerance behaviours and stigmatise low frustration tolerance behaviours, and, as a society with a shared, coherent belief in this, work collectively to encourage desirable behaviours when raising kids or teaching students, you aren't expecting kids to magically develop self-control by osmosis.

I think you could probably argue that successful reactionary societies fit this description: Confucianism, for example, is a high frustration tolerance superego code. There are also plenty of unsuccessful reactionary societies but those tend to allow elements of low frustration tolerance in, e.g. you have to submit your individual desires to the greater good of your kinship network most of the time but it's fine to beat your wife if she does anything vaguely irritating and also fine to kill your daughter if you find the embarrassment of her dating boys uncomfortable. And then there's the direction of travel in Western progressivism, which is that any frustration discomfort is problematic. I think it's this that Lasch is picking up on; it's just a shame he didn't read Ellis instead of Freud before writing *The Culture of Narcissism*.

# The Eldraeverse by Alistair Young

I will start off with a simple observation: things are not great.

I don't think it's really necessary to state specific examples; by and large, we see them on a daily basis. The world may be better than it has been in the past, but it still has heapings of misery and evil abounds.

And when faced with such things, us humans have a simple tendency; we imagine how things might be better. We create fiction, with all its myriad utopias and wonders. We invent civilizations where things are good, fill them with people kindhearted, and tell great quests where the good of the world triumphs over it's evils. And in doing so, we both create hope for a better future, and find ways to make it so.

But in my eyes, there is one critical detail that is all-too-often skipped over. The bad in the world did not come from nothing; the vast majority of humanity's woes can be blamed on humanity itself. World hunger is a solvable logistical problem, climate change would never have happened if people were willing to believe in it, and I hardly need to mention crime or war. The flaws of the world come from the flaws of the people in it, like a tiny crack growing to split an entire boulder.

This is where fiction often falls flat; they imagine a better world, but not better people. Oh sure, everyone we see might be kindhearted and generous, but evil is not the sum of humanity's flaws. Far more corrosive are the mental flaws, the biases and cognitive defects and just plain stupidity. Death itself could be delayed, frozen, and pushed back in near entirety if not for the fact that most people can only conceive of it as an inevitability, an immutable wall that cannot be broken, only accepted.

Fiction gives us Utopias, but rarely does it give Utopians.

But 'rarely' is not 'never'.

Perhaps the best example I have ever encountered is a little-known universe known as the Eldraeverse. Authored solely by Alistair Young, the Eldraeverse contains a modest three books and a blog. Written primarily as microfiction, the Eldraeverse depicts the vast and detailed Empire of the Star, a society spanning hundreds of lightyears. As work of hard science fiction, we see a culture that has grown and overcome the obstacles put in place by people and physics

both, creating a unique society where dozens of different species live together in what could well be called a utopia by some, or at least a weirdtopia.

And what of the utopians? That would be the titular eldrae, founders of the Empire. Where humans have all-too-often failed, they are driven to succeed. They experienced the pains and sorrow of a turbulent pre-history, and strived to overcome it. As a matter of longing for the worlds that fiction depicts, as a matter of disgust for the flaws of the world, and simply as a matter of *pride*, they tirelessly worked to create a better world. At first for themselves, and then for any others who would join them.

But what are the eldrae? Why have they succeeded where humanity fails? To put it simply, the eldrae are space elves. They are an offshoot of humanity, taken from a distant Earth by a species of precursor. Modified to suit the precursor's needs, they were relocated and then abandoned, as precursors are wont to do. This left the eldrae with a myriad of physical changes; blue blood and a pale complexion, biological immortality and slightly more resilience, and yes, pointy ears.

However, the physical changes are a footnote to the mental changes. Compared to humans, they are energetic and dynamic novelty-seekers. Lacking in what we would term 'childhood', they instead replace it with life-long wonder and enthusiasm. But at the same time, they're far from naive; their immortality has led to excellent long-term planning, they're nearly bereft of choice fatigue, and they have a will of iron.

Far from a perfect combination, in their early history the combination of intense emotion and determination led to no small amount of conflict. But from this conflict grew a unique ethos; an incredible respect for free will and self-determination, with an intense focus on obligation and individual responsibility. The eldrae are honest to an extreme, take their promises and agreements seriously enough to impress a Fae, and temper and direct their emotions with willpower and rationality that would fit in on Vulcan. They have an unending desire for perfection and awesomeness, both in themselves and what they create. And they take pride in it all; for what is pride but honesty about one's own excellence?

From this has grown a society that could perhaps be best described as libertarian; the right to self-determination expands into the right of owning that which you've earned, the respect for free will expands into the right to take on obligations, and much more. The end result is something like if you took a classical Fae, removed the excessive malice, and nudged them into creating a prosperous society.

And it *works*. The eldrae, in many senses, are not human. They don't think quite like humans, and they don't act like humans. They are, in the most literal sense, alien, with all the wildly different modes of thought that implies. And the eldraeverse takes this unique foundation, and expands it into what they would unhesitatingly call a utopia, and what we might as well.

In the modern era, the Empire of the Star is a vast civilization. Within it, countless alien cultures come together under a shared ethos. They live in cities of gleaming crystal and clean stone, travel in proud ships crossing the skies on plumes of fire. And everywhere they go, there are beautiful worlds to be found. And all these worlds are free. Free of pain, free of hunger and want. Entire worlds, free of death.

The Eldraeverse gives us a beautiful look into this carefully crafted universe. Far more than a mere glimpse, or a glossed-over society of hats. We see the dozens of species, ranging from vaguely humanoid to floating crystals and sophont solar turbulence. We see theramins played by octopuses and liquid-ammonia mixed drinks. We see clothes that are poured, houses that grow, and networks that sing. And, as a work of hard science fiction, nothing is left out; cornucopia machines are as common as coffee makers, micromachines are hard at work making sure your cape billows dramatically without a breeze, and the resident superintelligence rents their moon brains and plays matchmaker.

If you're a fan of science fiction, and you're tired of Grimdark, might I suggest some [Hopebright](#)?



# The Great American Gamble - Deterrence Theory and Practice from the Cold War to the Twenty-First Century by Keith B Payne

The Great American Gamble is a book analyzes how (ownership of) nuclear weapons was used and can be used in the future, to achieve important goals, like preventing nuclear war. It looks at the actual policies, but also the theory behind those policies. Were these policies and theories wise or were they [MAD](#)? Did the US avoid nuclear war by skill or by luck? Can these same policies keep the US and its allies safe from new threats, like North-Korea, China or terrorists?

## Theory - Schelling

The book contrasts two strategies to deter a nuclear attack, based on the ideas of two theorists, Thomas Schelling and Herman Kahn. Schelling argues that it is possible to achieve a stable balance of terror, where it is irrational for either side to launch a nuclear first strike, because the retaliatory strike will result in unacceptable damage. So no rational nation will then use nuclear weapons on the US or their allies, purely for reasons of survival. The biggest threat to this strategy is a loss of confidence in the retaliatory strike being sufficiently damaging to be a deterrent. This can be the result of first strike weapons that can destroy a large number of nuclear weapons of the opponent, as well as a partially effective ballistic missile defense system (BMD). In both cases, a first strike may be able to knock out the ability of the other side to retaliate with sufficient force, either because enough of their nuclear weapons have been destroyed, or because enough of their remaining missiles will be intercepted by the BMD. This can make a first strike a rational choice, because it removes the threat of destruction to your own nation.

Since a fully reliable BMD is (currently) impossible, Schelling argues that the goal should be to achieve a stable situation of mutual assured destruction, minimizing the fear on each side that their retaliation will not work, through unilateral and/or cooperative measures. Trying to gain an advantage by seeking first-strike capability or building a BMD is going to result in a dangerous arms race, which is Schelling calls an action-reaction dynamic, where the other nation will respond with increased offensive capability. So the rational choice is to avoid that situation, in favor of a stable situation of mutually assured destruction.

However, the question remains how the opponent is deterred from a smaller or

conventional attack? Is the American nuclear arsenal a deterrent to such an attack? Is it not preferable to be ruled by the Russians, rather than to die in a mutual nuclear exchange? The conventional military ability of USSR started to greatly exceed the capability of the West, so the Russians could also not be held back in Europe merely with conventional weapons. Schelling's answer to this was that the deterrent effect was not in the certainty of nuclear war, but in the uncertainty. It is not necessary to threaten with certain retaliation, but to make the credible claim that nuclear may happen due to reasons outside of the control of the one making the threat, like chance, accident, third-party influence, imperfect in the decision making machine or processes that we don't entirely understand.

To make this more credible, one can introduce complicating factors, like the presence of an American force in Germany. Furthermore, it is possible to dissuade opponents from making choices that may lead to nuclear war by "rocking the boat." This can be a bluff, but also a minor escalation. To be able to act this way, it must be possible to engage in partial attacks or retaliation, rather than merely through a full nuclear strike. Then any rational opponent would back down. Yet this is inconsistent, because it asserts that it is rational for the opponent to back down when doing otherwise would seriously increase the risk of nuclear war, but for oneself to escalate the situation, to dissuade the enemy. It ignores the possibility of a scenario where both sides consider their own demands to be reasonable, but the demands of the other side to be unreasonable, and expect the other side to back down if they escalate. Furthermore, it is also inconsistent for Schelling to recognize that nations may consider it very plausible for imperfect decision making to result in outcomes that the leaders didn't want, and be dissuaded by that possibility from taking actions that bring the nations close to nuclear war, yet not to see imperfect decision making as a way in which deterrence may fail.

Schelling gives the following analogy to explain why we don't have to be afraid while we depend on mutual deterrence, or have to hate our enemy to use it: "People regularly stand at the curb watching trucks, buses, and cars hurtle past at speeds that guarantee injury and threaten death if they so much as attempt to cross against the traffic. They are absolutely deterred. But there is no fear. They just know better." However, in reality, pedestrians do walk into traffic, because they are drunk, distracted by their phone or because they are suicidal. Nor do drivers always stay on the road, by accident or on purpose. According to the WHO, hundreds of thousands of pedestrians die every year, with many more injured. Car manufacturers consider pedestrian safety when they design their cars, even including active safety systems that merely benefit the pedestrian, for example, by raising part of the hood when an impact is imminent, to cushion the blow.

## Theory - Kahn

A competing theory is offered by Herman Kahn, who argues that a self-destructive threat on behalf of others is not credible. The Soviets cannot be expected to believe that the US will initiate a nuclear war to defend NATO allies, when the US will not survive the expected retaliation. Instead, the Russians will likely see it as a bluff. So the US can only credibly threaten nuclear escalation by being able to limit the damage from a retaliatory strike. He argues that being able to limit the loss to 10% of the US population might be sufficient to make the threat credible. Only the US needs to have the ability to credibly threaten with nuclear war, because the Soviets have superior conventional forces. Such survivability could be created by relocating people and industry, creating shelters, air defenses and a BMD.

Maximizing the credibility of the threats that are necessary to defend NATO allies requires a threat that leaves little to chance and that doesn't limit itself to merely targeting cities and civilian industry, as Schelling suggests. Instead, it should also be capable of striking at (hardened) military targets.

Kahn doesn't believe that deterrence is 100% reliable and that the suggested defenses that he suggests are not merely helpful to discourage Russian aggression, but also to reduce the damage if nuclear war does happen. Perhaps enough of society will remain standing to allow it to continue, rather than to end up in the Mad Max universe.

## Rationality?

Both Schelling's and Kahn's theories are ultimately based on the idea that there are universal goals, in particular societal survival, that motivate the leaders of nations. The problem these theorists are primarily concerned with, is the plausibility of the threat. Are nations sufficiently worried about their society being destroyed, to prevent them from launching a nuclear strike on the US, attacking Western allies that the US really doesn't want to be taken over by the Soviets, or take risks whose outcome they cannot oversee?

Yet neither theorist examines history to see if their definition of rationality is truly universal. Japan made a choice to go to war based on the belief that giving in to the American demands, or accepting the American boycott, would result in being relegated to a nation of minor importance. The advocates argued that this was an unacceptable loss, so a chance of victory was better than sure defeat. A prince of the Imperial family

wrote that if Japan lost: "there will be nothing to regret because she is doomed to collapse even without war." And at the end of the war, Japan preferred the enormous death and destruction of civilian resistance to an American invasion, with no hope of victory, over surrender, because they deemed a honorable loss more important than preserving life. Similarly, Nazi Germany preferred to fight a losing battle over surrender. Several of the top leaders were under no illusion about their fate upon surrender and preferred suicide. Would they have refused to use nuclear weapons out of fear of retaliation?

Both Mao and Fidel Castro expected an American invasion and one point, and suggested to Soviets to respond with a nuclear first strike from their territory (using Soviet nukes), rather than fight a conventional war. Both accepted an US nuclear retaliation on their nation and weren't willing to first see whether a win or stalemate could be achieved with conventional arms; or whether the goal of the invasion was limited.

However, it is not just the enemies of the US or those with a very different culture who may accept enormous destruction and death because they consider other values to be more important. The British decision to declare war on Nazi-Germany after the Polish invasion was assumed to result in mass bombings by the Germans on British cities, leading to hundreds of thousands or even millions of casualties in a few weeks. A post-war prime minister who served in the Churchill administration during the war, wrote that they considered the expected consequences of declaring war on Germany to be in the same category as later strategists considered the consequences of nuclear war. Yet they did declare war.

So it is highly doubtful that self-preservation or avoiding mass death and destruction is or will be the most important goal of all leaders who have nuclear capability. This is true in particular for leadership that has values or beliefs that are relatively alien to many in the West, like honor, destiny, martyrdom, reincarnation or a belief in an Utopia that is worth immense sacrifice to achieve.

Schelling recognizes the possibility that a leader may make a decision that he considers irrationally suicidal, but argues that those orders will not be followed. However, this ignores that orders that may seem irrational to Schelling, may not seem rational to people in another culture or with another ideology. Or, the military leadership may not be willing to question or countermand decisions.

## Deterrence during the Cold War

The actual policies followed by the US during the cold war were far more based on Schelling's theory than Kahn's. One of the strongest proponents and key figures in establishing US nuclear strategy for the next decades was Robert McNamara, who in 1962 said that he wanted the Soviets to have confidence that they could retaliate effectively. This is the logic of Schelling, where being vulnerable to the opponent and them being vulnerable to you, creates a stable situation.

The Pentagon calculated that offensive spending was far more cost effective than defensive spending. Reducing casualties would cost 3-4 times as much as the cost of new weapon capabilities to increase the casualties back to their previous levels. It was calculated that maintainable levels of spending would ensure at most a survivability of 80% of population and industry, which was considered insufficient. A loss of 20% of people and industry would destroy US society and thus would be utterly unacceptable. A senior White House official in the Kennedy and Johnson administrations argued that "10 million casualties are no different than 100 million casualties." By this he didn't literally mean that there was no difference, but rather, that a reduction from 100 million casualties to 10 million was not considered worthwhile. The National Security Advisor to Kennedy and Johnson even argued that only perfection was good enough for missile defense: "[nuclear weapons] make it necessary to defeat them all. Anything less is not good enough for safety."

McNamara assumed that the Soviets would have similar beliefs and being able to kill 20-30% of the Soviet population in retaliation would be sufficient to deter a Soviet attack on the US and its allies. He based the size of the nuclear missiles arsenal on that number. However, McNamara and other politicians seemed to diverge from Schelling's theory on the point of uncertainty. They generally seemed to commit to a retaliatory strike against the Russians and sought to minimize chance, for example, by trying to maximize order-following by missile operators.

Successive administrations and officials largely followed the same lines as McNamara, although probably because it was the most politically advantageous one. It was hard to sell large spending on defense, when those defenses didn't provide any obvious benefits. Promising 100% safety was also more appealing than Kahn's belief that safety could not be assured. The advocates of Schelling's balance of terror portrayed Kahn as a warmonger, who sought to start nuclear war. By Schelling's logic, his policies were likely going to lead to nuclear war, or at least increase the risk, but obviously that was not true according to Kahn's theory. However, it seems more likely that Kahn's insistence that we think through the consequences of nuclear war, is beyond most

people. The horror is so immense, that it is preferable to find a way to prevent it. A theory that claims to provide it, is far preferable to one that doesn't, even if significant criticisms can be made of the former.

However, merely recalling the Cuban missile crisis and the statement by one of the involved that the chance of nuclear war was about a third to even, should be sufficient to dismiss the idea that deterrence is highly reliable. And the decision by the Russians to 'destabilize' the situation by stationing nukes so close to the US, which makes a first strike more feasible, undermines the theory that the US can restrain others by restraining itself. The belief that one's own actions determine the actions of the opponent seems extremely narcissistic. While Schelling's theory asserts that there is a symmetric action-reaction dynamic, in practice the American politicians and advisers seem to always assume that the Americans act and the Russians react. The result is a strong inward-looking perspective, where the US is assumed to be safe if it manages to restrain itself. Fairly little recognition exists that the opponent may deal with the same difficulty to restrain itself (or have no desire to do so) and that this scenario requires a strategy as well.

The evidence strongly suggests that the Russians rejected the idea of stable deterrence, with no side having the upper hand. At least two studies analyzed American and Russian decision making and found no evidence that American choices to not develop certain nuclear weapons or BMD, restrained the Soviets. The Anti-Ballistic Missile Treaty, signed in 1972, states in it's preamble that restricting BMD would be a "substantial factor in curbing the race in strategic offensive arms." Yet it only caused Russian development of offensive nuclear weapons to accelerate, as the Russians could focus more on those. This outcome disproves Schelling's theory that rejecting BMD is (always) stabilizing.

While the Americans intentionally refused to build nuclear missiles that were both powerful and accurate, which could be used to attack and destroy hardened missile silo's, the Russians developed these and build a large number. It was only thanks to nuclear submarines that the US maintained survivable retaliatory capability. The Russians also built BMDs to defend Moscow, while the Americans didn't build one for their capital. Russia consistently appealed to the American theory by objecting to any American BMD that could defend against (some) Russian nukes as being destabilizing and a threat to peace, using language that is persuasive to adherents of Schelling's theory, yet they themselves didn't act as if they believed the theory. When Americans discussed their deterrence beliefs with the Russians, they rather consistently rejected the theory of the Americans.

The build-up of nuclear stockpiles peaked in 1986, shortly after Perestroika,

Gorbachev's attempt to restructure the Soviet Union to solve their poor economic growth. The stockpiles declined rapidly after that. The START I treaty that put limits on the stockpile, was only signed in 1991 after both sides had already begun to significantly reduce their stockpiles (between 1986 and 1991, the Soviets went from 40,000 to 29,000 nukes). Schelling had great faith in the ability for arms control through treaties, but in reality, it seems that these 'arms control' treaties merely tend to affirm choices already made.

## Deterrence after the Cold War

As the cold war progressed, deterrence strategy became more and more dogmatic, probably by the same mechanism that we see in many fields. The innovators think things through deeply and investigate different possible strategies, but later generations mimic their ancestors, but often without understanding. So they keep applying the same strategies even when circumstances change and invalidate the premises underlying the strategy. Furthermore, nuances get removed from the strategy that those earlier generations did advocate. For example, McNamara didn't advocate the same strategy against China as against the USSR, because he deemed it possible to build a BMD that could defend the US sufficiently from the limited nuclear arsenal of the Chinese, especially given their poor economy at the time. Even though it was just as true for China that evading American missile defense was much cheaper than building those defenses, China was so poor and technically incompetent that they couldn't maintain such an arms race. Yet after the cold war, a lot of strategists and politicians argued that BMD was useless and dangerous against any enemy, even the most weak ones.

In the post-war reality, there is a need for a recognition of the differences between opponents, including what kind of deterrence works on them (if any). For example, terrorists that value martyrdom are not deterred well by mass retaliation, if that is even possible, but seem to be deterred by being prevented from dying a 'true' martyr, by dying without being able to inflict casualties. So targeted strikes may deter others from choosing this path, or may persuade them to abandon it. On other opponents, nuclear deterrence may work well. Ultimately, it is important to not apply the same hammer to every nail, but to analyze each major opponent and choose a course of action that is likely to work best on them. It is also important to recognize that biological and chemical weapons are also WMD, whose use should be deterred.

Deterrence may require a different strategy or different weapons against different components. For example, a common occurrence during the 20th century was that opponents of the US or its allies, felt encouraged by supposed humanitarian 'softness'

that would make democratic nations unwilling to accept the deaths that were needed to stop their aggression. Examples of this are Hitler, Hideki Tojo, Mao, Saddam Hussein and Slobodan Milošević. This often led to lose-lose scenario's where those politicians made choices that they expected would not result in American/Western aggression and that they probably would not have made if they anticipated this. The more difficult to achieve solution to this, is to change the perception of humanitarian 'softness'.

However, US aggression can also seem more plausible if weapons can be used that are more surgical. The nukes that were developed for the cold war were high yield, low accuracy, perfect for destroying cities. However, it may be more deterring to many opponents if the US can use low yield, high accuracy weapons. For example, these weapons could then plausibly be used to threaten the lives of the leadership of those countries, even if they hide in well-protected bunkers, with little harm to civilians. Then the US can afford to be seen as humanitarian and very wary of collateral damage, while still deterring 'minor' transgressions, like the invasion of allied countries.

As the goals of the US and it's allies go beyond preventing nuclear war, those objectives should be considered as well. The book identifies four main methods for which nuclear capabilities can be used: deterrence, assurance, dissuasion and damage limitation.

Various allies have indicated that they place great trust in the US, including the 'nuclear umbrella.' Both Japan and Germany have indicated that they may acquire nuclear weapons if this trust is lost. So preventing nuclear proliferation among allies (which in turn may cause their opponents to acquire nukes), requires preserving this trust, in particular that a nuclear or existential attack on those countries will result in nuclear retaliation by the US. The reasonably aligned goals means that the allies can be asked about this, where their answer is likely to be much more trustworthy than when questioning opponents. For example, they may indicate that having American nukes stationed on their territory provides assures them that the US nukes will keep them safe. American nukes were and still are stationed in European nations, which may have contributed to Soviet reluctance to attack Western Europe, but also to confidence by Western nations in that their alliance with the US was worthwhile and that they didn't have to acquire nukes. Preventing nuclear proliferation among allies may prevent a cascading effect, where their (local) enemies also seek to acquire nukes. Preventing such a scenario may be one of the best ways for the US and their allies to reduce the risk of nuclear war, or of nuclear terrorism. After all, nations regularly descend into chaos, experiences coups or such, with government property falling in the hands of those who may not use it responsibly.

Dissuasion can also be used to prevent nuclear proliferation. For example, a BMD system that provides good defense against the number of nukes that a smaller, poorer or otherwise less capable nation can acquire, may convince those nations to not invest



in a nuclear program or even to abandon such a capability. Another goal of dissuasion can be to discourage China from getting into a nuclear arms race with the US.

Damage limitation is particularly useful against nations or actors with limited nuclear ability. BMD can be quite effective against a limited attack by nations like North Korea, terrorists, or rogue elements within powerful nations. In the case of a successful attack, many lives can be saved with relatively modest investments, for example, by stockpiling iodine pills.

## Verdict

I would definitely recommend the book. The public debate on this topic often appears dogmatic and simplistic, if it is not absent altogether. For example, advocates and critics of nuclear disarmament rarely bring up the possibility that nuclear proliferation may actually accelerate if allies lose their trust in the American nuclear umbrella, even though this seems very plausible. The value of nuclear weapons against contemporary enemies is also rarely debated, even though it is very important for people to understand the purpose of these weapons, lest they be abandoned without suitable alternative ways to achieve the same goals, or are misused. Trump allegedly asked an policy expert three times why the US couldn't use nuclear weapons, which suggests that the expert might not have been able to articulate this clearly. The reporting on this story seemed limited to attacking Trump for his foolishness, rather than actually explaining why they thought his question was wrong. Yet in itself, the question was fair. Government spending should be spent on things with a purpose and the purpose should be made clear to both politicians and the populace. I feel that the media are not doing their job on this topic. By reading this book, you can inform yourself reasonably well, or at least, as well as can be expected from one well-articulated and -sourced book from a specific point of view.

Yet that doesn't mean that the book is perfect. The author seems to occasionally exaggerate his sources. Insinuating that Fidel Castro wanted to unilaterally attack the US with nukes, rather than in response to an invasion, is misleading, for example. However, most of the book seems well-supported even if there are some exaggerations, because many of the claims in the book are fairly modest. The book primarily seeks to rebut the certainty of others and suggests instead to tailor the nuclear strategy to specific opponents (and allies), rather than assuming that the same strategy works for every opponent/ally and in any circumstance. Suggestions are offered, but never claimed to work with certainty, so readers are expected to draw their own conclusions.

What I found frustrating about the book is its structure. Instead of exhaustively exploring a topic and then moving on to the next, the book keeps revisiting subjects, where they keep getting expanded upon. Instead of continuing where it left off, a lot of the previous material is summarized. The book already extensively quotes politicians and advisers, even if they have largely the same beliefs as others, which does make it clear how nuclear strategy (didn't) develop over time, yet also results in a feeling of repetitiveness. Combined with the tendency to summarize earlier material, the book feels very padded. Fortunately, there is an index, so the book is still usable as a reference, as the structure makes it very hard to find anything otherwise.

# The Gulag Archipelago

A priest, a Communist and a farmer walk into a bar.

No, no they don't. But as dire as the Gulag was, there was certainly an interesting cast of characters. An Estonian cabinet minister might share a cell with a starving peasant who had stolen unharvested grain at night. A renowned Moscow engineer might haul a tree trunk with an illiterate who had accidentally desecrated a newspaper picture of Stalin while practicing how to write his name. A right-wing Ukrainian rebel might drink gruel beside a bewildered Communist. But whoever you were and whatever you'd done, or hadn't done, you were all prisoners. Together you would endure savage interrogations, stifling train journeys, frigid arctic winters, exhausting work assignments and bleak exile. And one by one you would die, of starvation or typhus or pellagra or scurvy or beating or execution.

Solzhenitsyn was a Soviet officer arrested in the last months of WWII. He was a Communist at that point but had some reservations about particular policies, and he exchanged letters with a friend in which he shared these views. Later when he recounted how he and his friend were caught, other inmates were amazed he would be so stupid as to criticize Stalin in a letter that would obviously be read by censors and pronounced him and his friend the two biggest idiots in the Soviet Union. He would barely survive his eight years in Gulag but the experience would transform his political and religious views, and would allow him to write one of the most important non-fiction books of the 20<sup>th</sup> century, *The Gulag Archipelago*.

The three volumes largely derive from Solzhenitsyn's memories, those of the numerous prisoners he spoke with, and the contents of letters sent to him from various ex-prisoners and officials during the less restrictive Khrushchev years. He orders Volumes 1 and 2 roughly along the lines of the experience of an individual convict: arrest, interrogation, trial (didn't happen for common prisoners, but of course Show Trials were a thing), sentencing, transit, and life in the labor camps. Volume 3 is about the changes and rebellions in the camps during the early 1950s, and Solzhenitsyn's experience and observations of exile in Kazakhstan. For a compendium of atrocities, it's quite readable. Solzhenitsyn combines powerful storytelling, scholarly analysis and scorching Russian humor.

## Central theses

Like any three-volume account of anything, there is a lot of ground covered over and over in *Gulag Archipelago*. Solzhenitsyn clearly wants to establish the following two points above all others: 1) The Soviet Union was far more brutal than Czarist Russia and was not in any conceivable sense a step up and 2) the brutality was not an aberration of the late 1930s, or even of Stalin, but started with Lenin and continued at a lower scale under Khrushchev.

On 1), there's no question that [executions were far rarer](#) in Russia under the Czar. They numbered in the hundreds during the entirety of the 19th century. There was an uptick in executions in the last 20 years of Czarism, but even then they occurred at a rate of about one hundred per year. I might quibble with the numbers. Pre-[emancipation](#), [nobles abused serfs](#) outside the purview of the formal legal system. Pro-czarist paramilitary [Black Hundreds](#) killed several thousand people in the revolution of 1905. Even the traditional Russian aversion to capital punishment for non-political crimes has an elitist edge to it. A nobleman might feel very Christian giving a light sentence to a bandit who murdered a couple serfs. Serfs might take a dimmer view of such leniency. If the English hangman was busier than the Russian one, perhaps that's precisely because England was more democratic. But whatever, Solzhenitsyn is right. The scale of executions under Communism was orders of magnitude worse than under the Czar.

Likewise, at least in the last fifty or so years, Czarist courts functioned as courts. People had rights, the courts sometimes [acquitted people the government didn't like](#). Communist trials were about showing off how bad the defendants were, not about allowing them a chance to prove their innocence. Prisons were also generally less awful under the czar. People usually got enough food and weren't tortured. Solzhenitsyn tells the story of an old activist whimpering after awful Soviet torture and remarking that under the Czar the guards weren't even allowed to address him impolitely. You [can find](#) instances of starvation and beatings among exiled prisoners during Czarism, but the scale was far smaller.

On 2) he is also correct. First there was a spate of mass killings during and immediately after the civil war, which was admittedly somewhat understandable in context. There was also a [famine](#) during this period of similar magnitude to the Ukrainian famine of the early 30s, but again it was partly the fault of "war" rather than of Communists so it gets less attention.

During the subsequent moderate period (~1922-1928) things were vaguely less horrible, but people who had been members of any non-Bolshevik political movement got arrested or stayed arrested in this period and were shuffled endlessly from Gulag to exile (being forced to live in an isolated small town on parole) back to Gulag. Religious believers and clergy were also arrested through the 1920s with little fanfare, and nasty tactics were used to extort hidden jewelry or gold from people suspected of hiding such goods.

The collectivizations of the early 30s swelled the gulags and the number of executions spiked during the Great Terror (1936-1938). The late 30s were when [famous old Bolsheviks](#) met bad fates, which allowed people not paying much attention to imagine that this era was an aberration. WWII meant some people from Gulag were allowed to fight, but also meant that large numbers of Vlasovites (Russians who fought for Germany), Ukrainian partisans and even Soviet soldiers who surrendered and were recaptured were sent to Gulag. Millions were interned until Stalin's death in the 1950s.

Under Khrushchev the vast majority of [political prisoners were freed](#), but even then thousands remained in custody and the state made a sneaky policy shift. Under Stalin, lots of essentially apolitical crimes were classified as political, whereas under Khrushchev political enemies were arrested under seemingly non-political grounds. Apparently they realized it just looked bad for the Soviet Union if lots of people were still sabotaging the revolution forty years after it happened, and so they would sentence political targets for crimes like "[hooliganism](#)" or even unrelated crimes like rape.

## Resisting interrogation

Solzhenitsyn's own words describe the mindset you need better than I can: "From the moment you go to prison you must put your cozy past firmly behind you. At the very threshold, you must say to yourself: 'My life is over, a little early to be sure but there's nothing to be done about it. I shall never return to freedom. I am condemned to die, now or a little later. But later on, in truth, it will be even harder so the sooner the better. I own nothing, I am dead to my loved ones, they are dead to me. From today on, my body is useless and alien to me. Only my spirit and my conscience remain precious and important to me.'"

Suffice it to say, this is very hard, especially the pretending your family is dead part. The few success stories were generally extremely religious types. The best most people could hope for was not incriminating others, and even then they usually failed under some combination of trickery and torture.

## Sentencing

One of the strangest part of the Soviet system was sentencing. Most political offenders (called "Politicals") got ten years in prison no matter what they did. This was true of people convicted of trivial crimes (expressing some sort of dissatisfaction with socialism, stealing small amounts of food from collective farms), trumped up nonsense (collaborating with the United Kingdom and Japan to sabotage steel production) or real rebels and dissidents (although these were often just shot). Superficially this makes no sense. Why doesn't the justice system distinguish between the guy who smirked at an article in Pravda and an armed Ukrainian rebel? It doesn't make sense until it does. You are either on our side or not, and the punishment for "not" is ten years in prison and a life as a parolee.

## Traitors

Traitors came in two varieties, imaginary and real. The imaginary were arrested *en masse* in the late 30s. The whole thing was ridiculous, hundreds of thousands of people arrested for spying for some combination of enemy countries. What's amazing in

hindsight is how most people believed this was really happening, including Soviet leadership.

The real were the hundreds of thousands of Russian soldiers who fought, or more often did non-combat work, for Nazi Germany or worked as officials in Nazi-occupied parts of the USSR. Since Solzhenitsyn was arrested near the end of the war, he interacted with a lot of these types. Their motives were mixed. On the one hand, many were just starving people who saw a way of getting out of prison camps where several million captured Soviet troops indeed starved. As Solzhenitsyn put it, "people who have been forced to gnaw on bats for sustenance have been released from all moral obligation to the Motherland, or the rest of humanity" (please remember this sentence was written pre-COVID). On the other hand, some people really hated Communism enough to fight for Germany.

At any rate, the military effectiveness of the German-Russian units was limited. The main thing the formal "Vlasov's army" did was prevent the Nazis from destroying Prague near the end of the war, and the Cossack and Tatar units were fairly small. Generally speaking, the Germans were too committed to the idea of Slav racial inferiority and enslavement to make much use of Russians until they were already losing the war, at which point nobody really wanted to fight for them. The "what if the Germans had been nice to Russians" alt-history is too stupid to even contemplate.

Arguably, the surprise is how many "Kulaks" and other victims of the Soviet Union were willing to fight for the motherland when it was invaded.

## Wreckers

Lawyers and judges of the old regime were of no use and had to go, but engineers were a different matter. You can't just execute all of your engineers, even the ones with bad politics. The result was that for twenty years the Soviet Union was dependent on people it mistrusted, until it could educate engineers from "friendly" classes.

This was not a sustainable equilibrium in such a paranoid state, and there were periodic "wrecking" trials for engineers. Wrecking was deliberate sabotage of production, and wrecking trials featured a mixture of the subjective and the absurd. The engineers took time to decide something, that was proof they were deliberately stalling to reduce Soviet production. The engineers built a high ceiling in a textile factory so that the workers wouldn't overheat on summer days: they were deliberately running up costs. The engineers bought less than maximally automated machinery from abroad because the Soviet Union was manpower rich and cash poor. That was proof they were buying obsolete equipment on purpose. Of course, if they had not made careful decisions, forced workers to sweat on summer days, and wasted money on super-advanced equipment, they could have been tried for those crimes.

Oh, also the engineers were part of a combined French/British/Romanian plot to invade the USSR. Sure, why not.

Wrecking became something of a national obsession, and people believed in it. Historian Simon Sebag Montefiore notes in *Court of the Red Czar* that industrial accidents were a natural consequence of breakneck industrialization, which made it easier to imagine sabotage everywhere. Solzhenitsyn remarks that even people put on trial for other nonsense crimes generally believed in wrecking, and "wreckers" likewise believed that there were spies and traitors everywhere. It took years in the Gulag to realize that it was all nonsense.

## When everyone is a criminal, the cop is king

The documentary "Street Fight", about Corey Booker's failed first run for Newark mayor, was an eye opener. If you put a Cory Booker sign on your business or hosted a Cory Booker event, the police would find a bunch of code violations at your business and shut you down. If you have enough codes, police have this power. The Soviet Union was this on steroids.

One of Solzhenitsyn's cell mates was a young engineer of good working-class origin who was in prison for a couple of seditious jokes he made. But not really. He was in prison because he refused to donate building supplies for a judge's vacation home, and so they dredged up the jokes from his file and threw him in prison. There's a number of stories like this. The secret police and judges seem to have been extremely aware of their power and used them to obtain material benefits and sex. They were helped by the fact that everyone was a potential prisoner. For example, if you survived behind German lines for any period and did any form of work, you'd committed a crime. They didn't arrest everyone in the western USSR of course, but they could have. As discussed above, an engineer who made any decision could be imprisoned for not making a different decision. And of course, they could always make stuff up.

## Labor camps

Labor camps were miserable. If you wanted to survive a ten-year term, you needed to avoid being put on "general work" (mining in a mining camp, cutting timber in a timber camp etc.) and get a special job. The prison barber wouldn't burn too many calories, the cook could steal food, and "trustees" were prisoners in leadership positions who could profit in various crooked ways. Not all trustees were bad, but few could resist opportunities for theft, even at the expense of the general prison population.

## The Saga of Naftali Frankel

Naftali Frankel stands out as one of the most remarkable figures in the history of Gulag, the great evil polymath. To this day his early biography is a mystery, but supposedly he was a merchant in the Ottoman Empire who made a fortune in the Black Sea timber trade. After the revolution, he was invited back to the Soviet Union to create a black market in gold. Apparently lots of people had gold hidden from the Communists which they were willing to sell for paper rubles but couldn't legally. Frankel was successful and was then arrested for black marketeering and sent to Gulag.

Through events which are also somewhat uncertain, Wikipedia claims he left an extremely detailed complaint note explaining how everything could be run more efficiently, he was made a trustee and then an official in Gulag. He pioneered the use of Gulag as a labor force and was allegedly responsible for creating a system that directly linked inmates' rations to their work outputs. Even upon rereading I don't quite understand it, but the gist is that workers and small teams of workers would be penalized if they failed to hit "work norms" but could get some extra food if they achieved X% above norms. It's worth stressing here that the people implementing this system called themselves Marxists.

In addition to boding poorly for weak or sick prisoners, Frankel's system had an unpredictable consequence. The extra food was never enough to compensate for the calories expended in exceeding work norms. People would burn themselves out trying to get six extra ounces of bread and die within the year. Prisoners would warn each other that "the big ration kills you, not the small one". This may have been a gross calibration error, or it may have been a feature. As with [slaves at Caribbean sugar plantations](#), it was cheaper to work people to death and get new people rather than to carefully keep people alive. Gulags never turned profits. The costs of guarding prisoners and their apathy to quality of work meant the best Frankel could do was loss-mitigation, so working people to death in a few productive months probably made economic sense.

Frankel himself seems to have been a genius. He was supposedly capable of remembering tens of thousands of faces and corresponding names and prison sentences, performing complex calculations in his head, and finding errors after glancing at architectural blueprints.

Solzhenitsyn has been accused of anti-Semitism, mostly for other stuff he wrote, but partly for his treatment of Frankel. In particular, he theorizes that Frankel was motivated by a hatred of Russians. Whether or not this is anti-Semitic, it is a silly argument. He's accusing Frankel of wrecking! This seems unlikely, both because wrecking isn't real and because there is an easy competing explanation. Frankel was extremely capable and completely amoral. If he'd lived in America, there would be university buildings named after him. Like the scientists, engineers, and artists extinguished at menial work, Frankel was a great mind wasted in gulag.



## Channeling Corruption

The last years of Czarism featured a bizarre combination of an effective secret police and a horribly ineffective prison system. The dreaded *Okhrana* could [insert informants deep into the ranks of the Communist party](#) and arrest top leaders. These leaders would then be sent to prison or Siberian “exile”, where they could communicate easily with other revolutionaries and pay a bribe whenever they felt like escaping. Stalin escaped from prison or exile eight times. It was mostly a matter of coughing up a couple hundred rubles.

The Communists, having first-hand experience with how bad Czarist prisons were at suppressing them, successfully created a system in which political opposition was completely squelched. They did so not by eliminating corruption, but by channeling it in directions that were harmless for the state.

The Thieves in Law (sometimes translated as “Brotherhood of Thieves”, hereafter abbreviated “Thieves”), perhaps best known in the west for their [coded tattoos](#), are an age-old presence in Russian criminality. In the prison camps, they effectively allied with guards and could terrorize inmates at will. Solzhenitsyn marvelled at how three or four Thieves could just take stuff from new prisoners one by one in prison cars filled with war veterans and nobody would fight back. He reflected that most people are only capable of courage if they feel they have societal backing, and that people correctly intuited that guards would back Thieves over Politicals.

The Soviet tolerance of Thieves was partly ideological and partly pragmatic. Despite its hardness, the Soviet state had some of the liberal idea that criminals were products of misfortune. Soviet law, at least initially, was very lax with criminals, producing a bizarre situation where you would get one year in Gulag for stealing all of a family’s possessions and ten years for stealing a bag of potatoes from the collective farm. The first crime was venal, the second ideological. On the pragmatic end, the Thieves were no threat to the Soviet state. Guards could indulge their corrupt impulses by profiting off trade with the Thieves (e.g. reselling stolen goods outside the camp at markups). Perhaps the Communists understood that corruption was inevitable, so it was better if it took the form of Politicals being extra abused rather than allowed to escape.

Periodically though, things would go badly for the Thieves. During the brief period when Yezhov was in charge of the NKVD (internal security ministry) and the number of executions skyrocketed, large numbers of Thieves were killed because of the perception that they were allies of the deposed NKVD chief Yagoda. There were fights between Thieves aligned with the guards (called “bitches”), culminating in the charmingly named “Bitch Wars”. Solzhenitsyn attributes the Bitch Wars to a kind of population control effort by warders but notes that he lacks the sources he’d need to write a chapter about them. Finally, around 1950, the Thieves were separated from Politicals for some reason, whereupon they no longer had anyone to prey upon.

## More corruption

The intersection of the Gulag and the broader Soviet economy/culture is its own interesting story. The word "tukhta" comes up over and over. It means something along the lines of "padding production numbers to fool the higher ups" and suffice it to say it happened a lot. Paroled prisoners ended up introducing the phrase to the Soviet lexicon, kind of the way gangster slang enters the American vernacular. The phrase "stretching the rubber", meaning "making a task take as long as possible because they will assign you a new task when it is done", also passed from the camps to the general population.

The best tukhta story involves a work foreman (himself a prisoner) who lied about the number of trees his crews cut down so that fewer men would work themselves to death. When the camp boss figured out what he was up to, the foreman made some argument to the effect of "you'll go to prison if this is every caught, so we're in the same boat now. Play along". And then the log-rafting office, the sawmill, and lumberyard didn't want to get in trouble so they dutifully processed the imaginary logs, which acquired additional attributes such as length, weight, construction-suitability etc. In fact, these enterprises were *happy* to process the imaginary logs; it allowed them to hit production targets without doing anything. And none of it mattered because the region had too much lumber to export so all the imaginary logs were allowed to float into the White Sea.

## Women

For a long time, many gulags were coed establishments, albeit with single-gender barracks. This had a number of consequences. The most obvious was guards, Thieves, and "trustees" taking sexual advantage of female prisoners. Again, this is the kind of corruption a state can tolerate. There were genuine romances too, and "camp marriages" that made life more tolerable. Some women seemed to have a powerful need to take care of a man, non-sexually. There are stories of starving old women doing domestic tasks for starving old men in terminal games of house.

At some point, separate female prisons were created. On the plus side, this provided some protection from sexual predation. On the negative side, women had previously been given easier jobs at camp, and when they got their own camps they had to chop down trees with the same work norms as men.

## Conservatives and kulaks

Conservatives like the Gulag Archipelago, and it's easy to see why. First, it bashes Communism. Second, it ridicules the idea of social engineering. Solzhenitsyn was interacting with the last generation who grew up under the czar and the first who grew up under Communism. In his telling, the latter were more lazy, cynical, selfish and venal. At the top there's the NKVD officer extorting sex from a woman who doesn't want to be arrested, and the bottom there's the free bakery worker bartering away bread

meant for prisoners on the black market. Whatever was supposed to happen, Communism completely failed to instill any sort of collectivist values in people. Third, the book is strongly pro-religion, and most of the people who successfully resisted interrogation or retained their moral bearing in gulag were driven by religious devotion.

The most interesting conservative argument pertains to the story of the Kulaks. Some quick economic history. During and after the Russian Revolution, land was redistributed and an extreme form of Communism (so called "War Communism") was imposed. Amid crashing economic production and mass famine, the Soviet state switched to New Economic Policy (NEP), in which the government controlled major enterprises but capitalism was allowed at the village level. Farmers grew food and sold it. Then c. 1929, the government embraced collectivization of agriculture and began killing, exiling and Gulaging so-called "Kulaks".

Kulak is often translated as "wealthy peasant" but if Solzhenitsyn is to be believed, there were no wealthy peasants. Everything had been made equal in the revolution, and disparities c. 1929 reflected hard work and ability from 1922-1928. "Kulak" was just the Soviet word for "above average farmer". Kulaks had to be found in every village, so you didn't need to have a commercial empire to be a kulak, just be a little better off than your neighbors. Some guy improved his house in the spare time, or cleared a few acres of wasteland for extra money? Die, Kulak. Meanwhile, worthless sorts made good "activists" and were happy to embrace the redistribution that came with collectivization. The lazy prospered, the capable were killed. People learned that individual initiative was dangerous, agricultural production sunk, and the moral character of the society rotted.

That's some Ayn Rand shit I just wrote, but it's real-life Ayn Rand shit, not Galt's Gulch. The hero isn't some mythical architect but rather a guy who ran a small sausage-making concern from 1924-1927 and died young in the arctic. And the victims are far more human. They grew potatoes, worshiped God, worked with their hands, and were starved, shot, exiled or worked to death for no good reason.

I should clarify that I am not now, nor have I ever been, a member of the Republican party. I also think it's a tacky business to try to pretend Joe Biden, or even Bernie Sanders, is somehow part Stalin. But reading Solzhenitsyn is a good reminder of the value of some conservative ideals. There's certainly something to be said for the doctrine of "economic freedom", the idea that people are better off if they don't rely on the state for survival. Ultimately, whatever ideology the state has, people in power use it for venal purposes, and it's not great if the same entity controls your benefits, the courts and everyone's job. Furthermore, hard work, self-reliance and ingenuity are virtues that need to be cultivated alongside compassion and empathy. Every society needs some Kulaks.

## Changes 1950-1954

Solzhenitsyn's life took something of an unusual turn. At some point, a survey was distributed to inmates asking about their skills. People made stuff up and he claimed to be a nuclear physicist. This wasn't totally *not* true; he was quite good at math and science. Anyway, on the basis of this he was taken to a "paradise island", a prison where he was fed normal amounts and given work in an office. Then around 1950, he was fired for spending too much time writing poems at work and sent back to the camps.

By this time, the camps seemed a bit less harsh. I suspect the real driver was that the Soviet Union was short on men post-WWII and realized it could no longer work them to death in months. But there were a couple factors Solzhenitsyn could observe. First, the removal of Thieves to separate camps. Second, the large-scale internment of [Ukrainian rebels](#) from the newly conquered parts of Poland. The camps could handle hapless Soviet citizens protesting their innocence but had a harder time with actual rebels. The Ukrainians were quite good at attaching electrical tape to sharpened metal spikes and murdering stool pigeons and devoted Communist prisoners. Other prisoners followed suit. In 1945, sharing the wrong thought in earshot of the wrong guy could get you another ten-year sentence. A few stabbings later, there was no more tattling and prisoners could converse freely. The death of Stalin and particularly the purging of ex-NKVD chief Lavrentiy Beria had another big effect. As the higher ups in Gulag got executed for their ties to Beria, the guards started to get very nervous and a lot nicer. To quote the CIA interrogator in *Zero Dark Thirty*, "you don't want to be the last one holding the dog collar".

In several prisons, there were outright rebellions in 1953/1954. These [didn't end well](#). Nonetheless, eventually, most prisoners were freed and given pardons of sorts, and the remaining prisons became markedly less hellish.

## Factuality

The main contemporary criticism of Gulag Archipelago was that, in the words of Solzhenitsyn's ex-wife, it's a bunch of "camp-fire stories". It's true that the book is largely a collection of anecdotes and Solzhenitsyn anticipates and pre-empts this criticism. Virtually all relevant archives were closed, anecdotes were the only evidence he had access to. If the Soviet State disputes his estimates of death tolls and mortality rates, he argued, they should open the books.

With the fall of the Soviet Union, more records have become available. There's clearly a consensus that Gulags were terrible places and I've seen no evidence that Solzhenitsyn was personally dishonest. That said, the death tolls he estimates at various camps and projects are generally considered too high, and there's extensive dispute over the mortality rates in Gulags.

[Wikipedia cites estimates](#) that 18 million people shuttled through the gulag and only 1.5-1.7 million died, plus another ~800,000 outright executions. *The Gulag Archipelago*

gives a strong impression that a ten-year term on "general work" was usually a death sentence, and that certain assignments invariably killed you in months. Lots of people did short stints for things like negligence at work, but the numbers above are just too low to fully support Solzhenitsyn's assessment. Other scholars cite far higher numbers (even Russian archives can suffer from tukhta) and there seems to be some dispute about how deaths in work sites vs. more permanent camps vs. executions are counted.

Occasionally Soviet apologists cite the lower numbers as proof the Soviet Union was a perfectly nice place. All I can say is that "if you told a joke about Stalin and got sent to Gulag for ten years, there's a decent chance you would survive" is not a very strong defense of a political system.

## Ending

The novel *Red Plenty*, which [Scott reviewed on the old blog](#), ends with fictionalized Khrushchev contemplating life immediately after being deposed from power. He thinks about all the bloodshed he has seen, all the people he helped kill during the Stalin years. He could rationalize these things because he reasoned that he was building a superior economic system, one that would deliver so much joy and prosperity that a few hundred thousand people shot in 1937 would be a rounding error in the grand scheme of things. With the ascent of Brezhnev, he realizes that will never happen. All those people died to produce an inefficient, mediocre, cynical bureaucracy.

The first big Gulag labor project was the [Belomorkanal](#), a canal which connected the Baltic and White Seas. Built with minimal equipment and a complete reliance on convict labor, the Belomorkanal was celebrated throughout the USSR. A collective of writers, led by Leopold Averbakh and the world-famous Maxim Gorky, published a series of essays describing it as a humane project that restored prisoner's souls. Not a few of these writers, including Averbakh and possibly Gorky, would meet their own fates at the hands of the security state they extolled.

Some years after his freedom and post-Stalin rehabilitation, Solzhenitsyn visited the Belomorkanal and learned its darkest secret. Not all the dead prisoners, everyone knew that. It's just a crappy canal, too shallow and shoddily built to support most economic activity or be of much military use during WWII. The day he visited, he saw exactly two ships. Each was carrying what appeared to be firewood, and they were travelling in opposite directions. Illiterates and intellectuals, monks and Tolstoyans, priests, Communists and farmers had toiled until they collapsed in the snow so that some firewood could drift up and down a pointless canal. That's the story of Gulag.

# The Horse, the Wheel, and Language by David W. Anthony (2007)

Indo-European is by far the most successful language family. 10 of the top 15 most commonly spoken languages are Indo-European:

English, (Chinese), Hindi, Spanish, French, (Arabic), Bengali, Russian, Portuguese, (Indonesian), Urdu, German, (Japanese), (Swahili), Marathi

This is not just a result of European colonialism: through most of history, including today, the largest Indo-European state has been centered in either India or Persia.

Who were the original Indo-Europeans? How did they come to such dominance?

To answer this question requires both linguistics and archaeology.

By comparing widely dispersed, ancient and modern, Indo-European languages, we can reconstruct Proto-Indo-European words. Some of these recreated words refer to ideas that can be found in any group: pronouns, body parts, and family relationships. Some of these recreated words refer to climate: snow and lynx, but not tiger and elephant. Some of these recreated words refer to technology: bread, cheese, wool, mead, axle, and multiple words for wheel. Some of these recreated words refer to aspects of their society: warrior, priest, oath, and guest/host.

We can then look at the archeological record to determine where the Proto-Indo-European people lived.

This history is not yet settled science, but agreement is growing along multiple lines of evidence. Anthony has been in the middle of these debates. This book contains some of his original work.

Proto-Indo-European is associated with the domestication of the horse and the advent of wagon-based nomads in the steppes of Ukraine and southern Russia from 4000-2500 BC.

Farming and domesticated animals had spread from the Middle East to the Lower Danube Valley (Old Europe) by 6000 BC. From there, domesticated cattle, sheep, and goats spread onto the steppes. Limited agriculture also spread to river valleys. Although domesticated animals allow more productivity, they also involve changes to society. Herds allow greater wealth concentration and result in more frequent raiding. People on the steppes west of the Urals adopted domesticated animals, but the people east of the Urals remained hunter-gatherers.

Many of the cultural features of the Proto-Indo-Europeans developed during this time. Large, raised mounds containing the graves of chieftains or warriors (kurgans), containing black, red, and white dirt, parts of sacrificed animals, and weapons, including animal-head stone maces. Funeral feasts that could feed hundreds accompanied some of the larger kurgans. Themes common to later Indo-European religions may have also begun: Sky Father & Earth Mother, a primordial cow, division of society into priests-warriors-commoners, epic poems, and oaths of fealty.

Horses were likely domesticated on the steppes by 4000 BC. They were likely originally raised as food that requires less winter maintenance, but soon were ridden as well. Anthony's own research has shown bit wear on horse teeth by 3500 BC. War was revolutionized by the increased mobility of fighters, not by the use of horses in battle.

The first migration off of the steppes into Old Europe also began. Most of the cities in the Danube Valley were abandoned, due to climate shifts and increased warfare. The evidence for the migration is primarily from kurgans in the Danube delta, so we don't know whether it was a mass migration or an elite migration that became culturally dominant. These people would probably go on to form the Hittite Empire in Anatolia, the first written Indo-European language.

After a few hundred years, Old Europe began to reassert itself, building the largest urban areas that had yet existed (~10,000 people). Unlike Ur or Sumer in Mesopotamia, these were towns, not cities. There is no evidence of political hierarchy or centralized power. About every 70 years, the town was burned and a new one founded at a nearby site.

Across the Black Sea, the northern slopes of the Caucasus Mountains come in contact with Mesopotamia. Kurgans start having finely decorated silver, complete with Mesopotamian symbols of royalty - the lion & bull. Wheels spread north onto the steppe.

Between 3300-3000 BC, a single culture spread across the entire European steppes, the Yamnaya Horizon. Kurgans appeared for the first time days away from water sources. Some Yamnaya kurgans include wagons. Wagons allow people to become nomadic in the steppes and to journey, with their herds, for days across the dry grasslands. Previously, only areas near river valleys had been inhabited. Now, the entire steppe could be used. The Yamnaya Horizon marks the group who spoke Proto-Indo-European.

This wildly successful group expanded more. First, a group crossed Kazakhstan, past the horse-riding hunter-gatherers. The descendants of this group would become the Tocharians, who lived in Western China until 800 AD. Then, they expanded across Europe. The ancestors of Celtic and Italic moved up the Danube River, the ancestors of Germanic moved west north of the Carpathian Mountains, and the ancestors of Slavic and Baltic moved north into the forest. The Corded Ware Horizon marks the expansion of Indo-Europeans across northern Europe.

These migrations, and the other Indo-European migrations before and after, are more likely elite migrations than mass migrations. Imitation of wealthy immigrants and assimilation of ambitious

individuals likely played a larger role than conquest and domination. Successful herders could amass huge amounts of highly mobile wealth. Chiefs would migrate to find new clients, and then established their ritual and social institution in the new land. Oaths of fealty allowed strangers to fully join their society. Indo-European religion demanded that the right rituals be performed in the right language, but was not concerned with ancestry.

At around 2000 BC, more changes are occurring at the southern end of the Ural Mountains, the eastern edge of the Proto-Indo-European homeland. Metallurgy became extremely common - there are heavily fortified settlements within which every house contained tools for smelting bronze. Much of this bronze was traded south to the cities of Central Asia, then through Persia to Mesopotamia or India.

The earliest chariots were found in their kurgans. Unlike wagons, chariots are designed to be light and fast. Their wheels are spoked, not solid. Chariots revolutionized warfare. They were the first horses to be used in battle - true cavalry wouldn't be developed for another thousand years.

Chariot warriors spread south across Central Asia to already-old civilizations of Persia, Mesopotamia, and India. These civilizations either rapidly adopted chariots or were conquered. They brought with them early Sanskrit and Persian, and enter the historical record.

The Vedas, the oldest Hindu texts, describe the Aryans entering India from the north, before 1500 BC. They contain funeral hymns that hearken to the kurgan burials and feasts on the steppes. In Syria, the Mitanni Kingdom had Sanskrit-speaking elites who worshiped Hindu gods (contemporary with Moses). The Mycenaeans presumably came around the other side of the Black Sea to establish the first truly Greek civilization. Their society was described in the Iliad and Odyssey.

From here, we can follow the historical record, rather than relying on archaeology and linguistics.

The story I've told here is much simpler than the one told in Anthony's book. I have purposefully neglected to mention the names of the many archaeological cultures described (including Poltavka, Potapovka, Pokrovka, and Petrovka). My story also has no pottery - although there are pages of pictures of pots. Anthony is an archaeologist. The archaeological detail does make the book harder to follow. Luckily, each chapter has at least one map showing the cultures and sites described in the chapter. When reading it, I kept a finger on that chapter's map and another finger in the endnotes. Anthony is trying to tell the entire history of multiple groups of people for thousands of years, so a complicated story is expected. The book contains a remarkable amount of information.



One linguistics question I have after reading this book:

When describing how to figure out Proto-Indo-European root words from modern and ancient languages, the book stated that certain sound shifts are more likely to happen than others. For example, hard consonants (k,g) are more likely to shift to soft consonants (s,sh) than vice versa. Then where did the hard consonants initially come from? There are other ways in which languages tend to shift in certain directions. How did proto-languages get to such a linguistically unlikely state to begin with?

There is a similar not-fully-solved question in physics. Since we know that entropy always increases, the early universe must have had very little. How did the early universe get into such a low entropy state?

Other similar problems in physics have been solved. If uranium (and various radioactive isotopes) decay, where did the original uranium come from? The answer is from events that release extraordinary amounts of energy (supernova & neutron star collisions), a small fraction of which is used to make radioactive elements and fling them across the universe.

Is there any research on this entropy problem of linguistics?

I can think of several possibilities:

- (1) Although the various linguistic processes have direction individually, collectively they form cycles. A possible example I just made up: hard consonant -> soft consonant, soft consonant -> vowel, too many vowels -> glottal stop, glottal stop -> hard consonant.
- (2) There was far more linguistic diversity in the stone age than today. [It's not clear to me if this would be true. Less travel suggests more languages, but fewer people suggests fewer languages.] Only a few of these languages have descendants. The ones that survived happened to be ones that were linguistically unlikely. This also suggests that there was some selection bias here. Perhaps some languages made people more likely to adopt agriculture or new technology and so outcompete their neighbors.
- (3) Linguistic rules are different in nomadic societies than in sedentary societies. Or some other division might be important. If this were true, it would have modern implications.
- (4) Divine creation of language, which has subsequently decayed.

# The Human Predicament: A Candid Guide to Life's Biggest Question by David Benatar

## 0. The Warning

*The Human Predicament* is a book that could ruin your life. Maybe this sounds like an exaggeration, but the author has [talked about](#) hearing from devastated readers who, among other things, now regret their children as “a terrible mistake.” “They have an accurate view of reality,” he says of these readers, “and they’re paying the price for it.”

The book asks hard questions — questions like whether life is worth living, or whether suicide is justified — that I think are worth taking seriously. But at the same time I acknowledge that there are many people who consider these to be mental health questions and not philosophical ones. If you’re one of those people you would probably be better off skipping this post and reading something more relaxing instead, like [this article on why we know so little about eels](#). I won’t take it personally.

If I had never read this book I might have remained a normal, well-adjusted person. But I can’t unread it so I feel I have a duty to provide a coherent response. Life *is* worth living, and it’s worth creating too, but to get there we need to talk about all the good reasons for thinking it might not be.

## I. The Asymmetry Between Pain and Pleasure

If the philosopher David Benatar is right, then human life is substantially worse than we think it is and there is nothing more important we can do than to avoid bringing more people into existence.

The bad news is that his argument is persuasive.

The basic problem is what Benatar calls the asymmetry between pain and pleasure. Pain and pleasure are ethical opposites but they don’t have equal weight. Benatar argues that this causes all kinds of moral conundrums that make it so that life is not worth creating.

To illustrate the point, imagine you and your husband plan to have a child. You go to your local soothsayer to receive a blessing. The soothsayer regrettably informs you that your child will be cursed to a life of miserable pain and premature death.

If the soothsayer is right, do you have an ethical duty not to conceive? Most people would instinctively say that you do. It’s bad to create pain in the world, and you do good by avoiding it.

Now imagine the soothsayer tells you that your child will live a comfortable life, will find meaning in their everyday routine, and will be loved by a good number of people before dying at an actuarially predictable age. But when you think about it some more, you realize that the business of child rearing — changing diapers, dealing with temper tantrums, losing touch with your childless friends — isn't all that appealing to you. You're starting to get cold feet.

In this case, do you have an ethical duty to conceive in spite of your doubts? Most people would instinctively say that you do not. It's good to create pleasure in the world, but it isn't bad to avoid it.

This, at base, is the problem: pain is bad, and its absence is good. Pleasure is good, but its absence is not bad. It is generally seen as reasonable to mourn the fact that suffering persists around the world, while it is considered unreasonable to mourn the fact that the universe is largely empty when instead it could be stuffed to the brim with conscious particles experiencing intense pleasure at every moment. If you care about maximizing pleasure and minimizing pain, Benatar argues, then each time you create a new life you're making a bet with bad odds. In the long run, the house always wins.

Unfortunately for us, the news gets worse from here.

## II. The Abysmal Quality of Life

According to Benatar, life is much worse than we think it is. It is "permeated by badness" (71) in big ways and in small ways.

The big ways are familiar: war, famine, pestilence, plague. Thomas Hobbes famously described life in the state of nature as "nasty, brutish, and short". This remains the case for most animals and for a lot of people around the world, although many of us in developed countries are increasingly insulated from this reality.

But even during the normal course of middle-class life in the United States, horrible things are always happening to us. We age incessantly, even though we crave youth. Our bodies are constantly under attack by diseases that only get worse as we get older. Most of our desires go unfulfilled — for sex, for wealth, for acclaim. Even those lucky enough to live the very best life anyone can ask for will wind up living long enough to watch both of their parents die.

Most of us acknowledge these major pains as being inescapable facts, even if we prefer not to think about them. Benatar's innovation is to argue that life is also horrible in lots of small ways that we don't recognize. He narrates the pain of everyday life:

"Even in good health, much of every day is spent in discomfort. Within hours, we become thirsty and hungry. [...] When we can access food and beverage and thus succeed in warding off hunger and thirst for a while, we then come to feel the discomfort of distended bladders and bowels. Sometimes, relief can be obtained relatively easily,

but on other occasions, the opportunity for (dignified) relief is not as forthcoming as we would like. We also spend much of our time in thermal discomfort — feeling either too hot or too cold. Unless one naps at the first sign of weariness, one spends quite a bit of the day feeling tired. Indeed, many people wake up tired and spend the day in that state.” (71)

Maybe you think that these problems would be better categorized as annoyances rather than pains. Fair enough, but the basic point stands: being alive is a physically irritating experience. To make things worse, these irritations are a feature of our biology, not a bug. Dissatisfaction provides a motivation to act, and so hunger returns every few hours to prompt us to seek more energy. These motivators, while useful for biological survival, are typically unpleasant.

Crucially, the discomforts of life are also *more* unpleasant than their relief is pleasant. This is the second asymmetry between pain and pleasure: pleasure is always short-lived, but pain lasts. As Benatar puts it, “chronic pain is rampant, but there is no such thing as chronic pleasure” (77). We get hungry an hour before lunchtime but the pleasure of eating a sandwich only lasts for a few minutes. (And on the other side of the sandwich is often a farm animal that suffers for hours if not days for our few minutes of enjoyment.)

If life is really all that unpleasant, why is it not the only thing anyone ever talks about? Well, to some extent it is. A lot of everyday conversation revolves around petty grievances and complaints, like bad bosses and recurring migraines. But most normal people don’t think of it as a major philosophical concern that they have to pee every few hours.

The key problem, Benatar points out, is that optimism bias pervades our experience, meaning that we’re terrible at making objective assessments of how good our lives are. We habituate to most changes to our circumstances such that amputees and lottery winners both revert to a baseline level of happiness after a short adjustment period. We judge our lives in comparison to others, rather than according to objective criteria of well-being. And we de-emphasize pain during memory formation, leading us to look back on horrible experiences like low-rise jeans and the 1970s with nostalgia rather than regret. This [hedonic treadmill](#) prevents us from understanding how uncomfortable everyday life really is and keeps us hopeful for lasting relief that will never come.

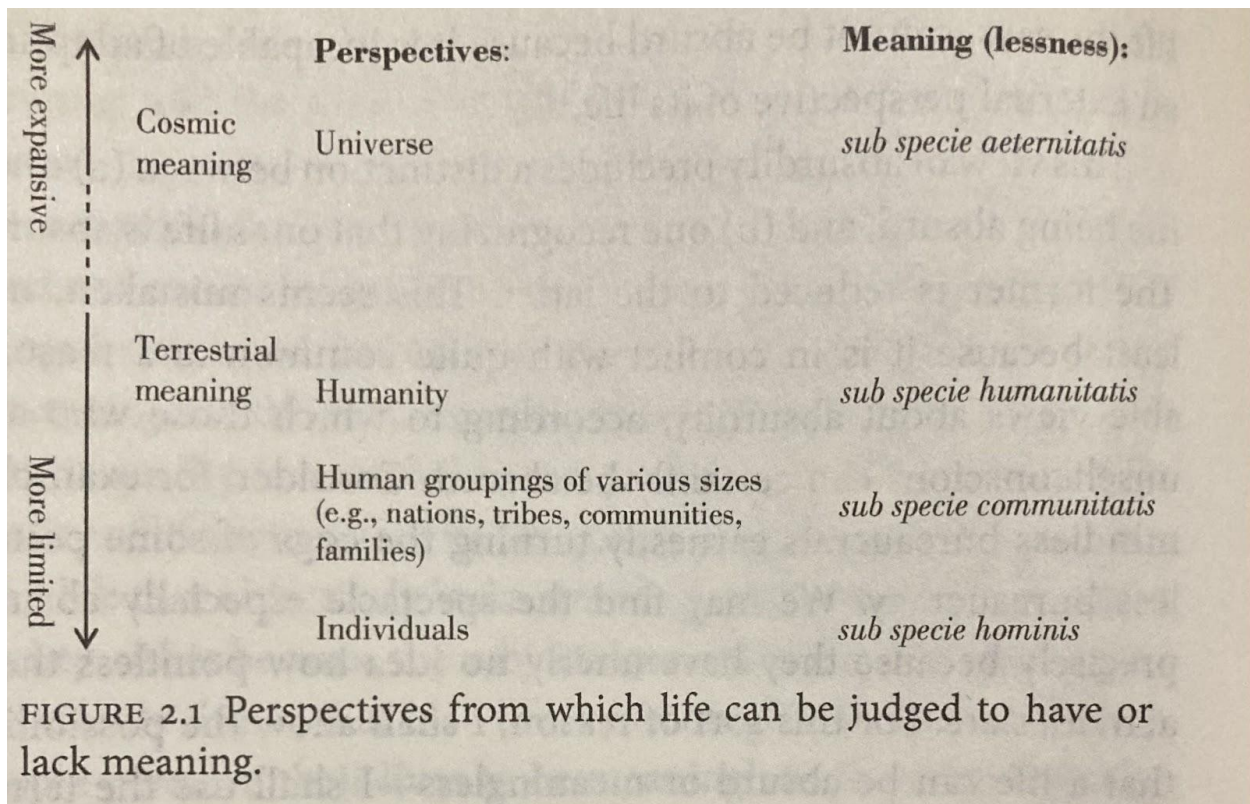
So life is full of pains that are worse, more ethically concerning, and longer-lasting than its pleasures, while we are built to misjudge precisely how bad the situation is. Surely this must be where the bad news ends.

### III. The Meaninglessness of Existence

Nihilism gets a bad rap. Normal people often sketch a caricature of nihilists as dorm room pedants, violent extremists, or [bumbling criminals](#). This reaction is understandable. Meaning is essential to people’s subjective well-being so it makes sense that it would seem both funny and a little bit threatening to suggest that life might not have any.

Philosophers sometimes skirt the issue by making complicated arguments about what exactly we mean by the word “meaning”. If this kind of question interests you, Benatar spends some time on an engaging summary of the debate. But he also points out that these arguments are only interesting to academics because most people share a common-sense understanding of what “meaning” means. For life to have “meaning” is to say that it serves some purpose beyond its own reproduction, that there are good reasons for something to exist instead of nothing.

Benatar does not espouse total nihilism. He acknowledges that there are some perspectives from which life can meet the standard of meaning. There are different levels of abstraction at which meaning can be defined, after all, and people are not wrong to believe that their lives have meaning on some of these levels. On the smallest level is the individual, then the community, ascending up to the human species, and finally reaching the cosmic level — a scale of meaning that would encompass the interests of the entire universe. In case you need help visualizing this, Benatar provides a handy chart.



Meaning is readily attainable on the smaller scales. The daily activities of going to work, preparing food and drink, and caring for relatives can serve clear purposes for individuals and their families, although Benatar points out that some people still fail to achieve this level of meaning in their lives and the experience of that failure can be awful.

As we ascend the ladder of abstraction, however, meaning becomes harder and harder to achieve. Many people can run for their local school board and find meaning at the community

level, but there are only so many seats and more people will lose than win. This dynamic gets worse as the degree of abstraction increases and there is more competition for fewer positions. Worse still, those who do achieve meaning at the higher levels often wind up doing so by causing enormous harm. Few have achieved the degree of impact that Stalin and Mao have.

At the very highest level — the cosmic scale — meaning is not attainable by anyone. This is typically the scale nihilists are talking about when they declare that life has no meaning: our small, short lives simply have no effect on the universe at large. As Benatar puts it, “even the most expansive terrestrial meaning will eventually vanish [...] if only because all humans will eventually become extinct” (95). It may be hard to imagine now, but even Donald Trump will someday be forgotten.

So meaning is more achievable at the lower levels of abstraction but it is also less satisfying. At higher levels it becomes more satisfying but less achievable, and sometimes dangerous. The problem is that the higher levels of meaning are the ones that capture our attention. When people struggle with “the meaning of life” they are usually not worried that their life serves no purpose to their immediate family members and coworkers. A crappy desk job can feel spiritually crushing even though you might admit that efficient payroll processing helps you deliver value to clients and shareholders.

Some suggest that the answer to the problem of decreasing feasibility of meaning at scale is that people should focus on cultivating the lower levels of meaning that they can achieve, a strategy sometimes known as [downshifting](#). This can indeed be a powerful coping mechanism and I recommend it to anyone who simply wants to stop worrying about this stuff. But from a philosophical perspective downshifting is ultimately a way of avoiding the problem, and once again our biology intervenes to make it increasingly difficult.

As with pain, our experience of meaning is also subject to a process of habituation. It can be exciting to land a new, more influential job, or to get a promotion that gives you the ability to manage a larger team. But spend a few months at it and you’ll probably realize that it feels just as pointless as your last job, only now there are more people asking for your time. Our brains encourage us to grow dissatisfied with the meaning we have in order to seek higher levels of meaning, which we eventually grow dissatisfied with, too. The cycle repeats, again and again, until eventually we reach our end.

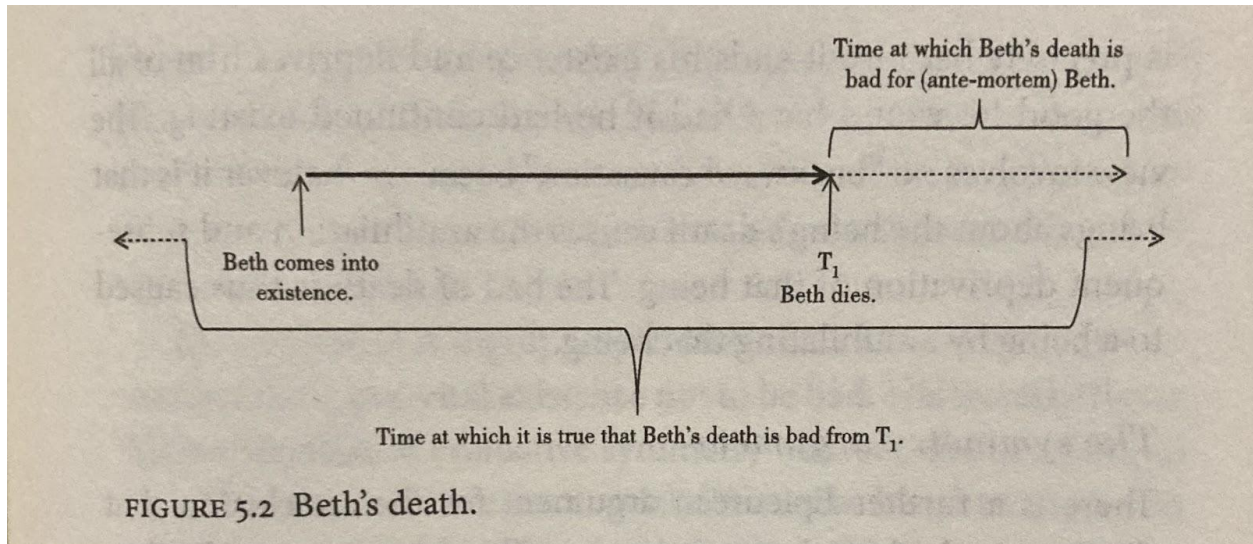
And this ending poses its own problems.

## IV. The Certainty of Death

Benjamin Franklin famously quipped that “in this world nothing can be said to be certain, except death and taxes”. ([Franklin didn't actually invent the idiom](#), but he did popularize it.) Benatar points out that this isn’t quite accurate since a variety of loopholes exist to help the rich cheat on their taxes while no living thing, rich or poor, can cheat death.



Like meaning, philosophers have found enjoyment in the semantic puzzle of whether and why death is bad. Epicurus first pointed out that death can't be that bad, separate from the suffering that often accompanies it, because the person who dies ceases to exist after their death, which means there's no longer a person for death to be bad *for*. Benatar provides another compelling summary of this debate, featuring fun questions like "When is death bad for the person who dies?" (110) and a chart illustrating the points in time at which you might reasonably describe your death as being a bad thing.



But here again Benatar argues that the debate over the badness of death is academic since most people share an intuitive understanding of it. Death is bad, Benatar says, because it *annihilates* us. It erases our particular perspective from the universe with no possibility of ever coming back. Sure, some fragments of us live on in the people who were influenced by our thoughts and actions. But eventually those people will be annihilated too, because death is the certain fate of all living things.

Not only does death annihilate us, usually causing us to suffer in the process, but it does harm to the people who love us. The longer we live, the more attachments we grow to the world, and the worse it is for everyone involved when we're forced to leave.

## V. The Predicament

This leads us to the "human predicament" that Benatar refers to in the title of the book: "Living is bad, but dying is worse". This is a predicament in the technical sense of the term, meaning that there is no easy solution to the problem. You can't solve the problem of life by dying, and you can't solve the problem of death by living. The whole thing is one big trap.

Hence Benatar's moral imperative to anti-natalism. Once a being is introduced into the world, it simultaneously suffers while also developing a strong interest in continuing to live. A being that never exists in the first place, on the other hand, doesn't have any particular interest either way. The only way to win the game is to refuse to play.

Benatar goes deep on a number of big questions related to the predicament. Is suicide justified? (For most people no, because death doesn't solve the predicament, but in some extreme cases that involve intense pain with no hope of recovery it can be palliative and we should take these cases more seriously.) Would immortality be good? (It wouldn't solve the predicament on its own, but 75 years is almost certainly not the optimal lifespan and we should strive to live longer if we can do it without degrading our quality of life.)

These are interesting questions but they are ultimately tangential to the main point. Benatar has made a strong case that life is bad, that death is worse, and that we are trapped by biology to perpetuate the cycle of suffering. He concludes that the decision that most people think of as the most selfless and important choice of their lives — the choice to have children — is one that we are ethically bound to avoid.

Perhaps you are fully persuaded by Benatar and you're adjusting your life goals to plan for voluntary human extinction. Or perhaps you don't find my summary to be logically sound, in which case I have probably failed to capture Benatar's reasoning. But if you think this argument makes sense and you have any sort of attachment to the idea that humanity should continue to exist, then Benatar has put the onus on you.

How are you going to solve the predicament?

## VI. The Objections

There are a number of standard attempts to refute Benatar's position. He anticipates many of them in the book, although some are better clarified than others.

One common response to Benatar is to argue that the pain of life gives clarity to its pleasure. By a related line of reasoning, death isn't all that bad, because it provides definition for life. Benatar characterizes this argument as the position that "bad things in life are necessary in order to appreciate the good things" (84). Unfortunately, it fails to engage with the asymmetry between pain and pleasure: most suffering in life is disproportionate, or is wasted entirely. Spending a few weeks apart from your boyfriend might encourage you to appreciate his company, but his death is unlikely to deepen your relationship; instead it will probably lead you to eventually forget about him.

Another common response is theological. The predicament is only a problem for atheists, this line of thinking goes, because God provides a divine plan that gives purpose to earthly suffering. Setting aside the question of whether we have reason to believe that a purposeful God exists,



Benatar points out that this argument sidesteps the problem of meaning. God's plan is fundamentally mysterious, after all, and just because He has a purpose for us doesn't mean that purpose is necessarily *good*. We have no way of evaluating what that purpose is, and Benatar argues that the few indications we do have aren't promising:

"If, for example, we are told that our purpose is to love God and serve him, we might reasonably ask why a being as great as God is said to be would possibly want or need the love and service of humans at all [...] If loving and serving God is our purpose, the act of creating us sounds like that of a supremely narcissistic rather than a supremely beneficent being." (38)

Still a third response, perhaps the most frustrating to me personally, is to agree with Benatar on the conclusion but not on the reasoning. Some people admit to me that they agree that having children is unethical, but they think it's unethical because children lead to a kind of Malthusian trap that is destroying the planet through climate change. This is wrong for so many reasons that it deserves its own post, but for now suffice it to say that the [carrying capacity of the planet is not fixed](#), that we need more people to produce the technological progress that will reduce emissions without lowering our quality of life, and that [climate change is unlikely to be our most pressing existential threat anyway](#).

Above all, the most common response that I get when I run this argument by people I trust is simple disbelief that life is really as bad as Benatar says it is. Life can be painful, sure, but if we can cope with the pain then we can see that it also involves fun things like snowboarding and tennis and video games. Death is bad, but there's not much we can do about it. So what's the big deal?

This position, I think, fails to take our optimism bias seriously. Suffering is not merely a temporary inconvenience on the way to a better life, but rather a fundamental strategy by which our biology urges us to survive and reproduce.

We are trapped by natural selection in a planetary competition, fighting for resources at the expense of other sentient beings, for no particular reason other than the reproduction of the competition itself. We can reduce the harm that we do and that is done to us in the process but we can't eliminate it because harm is built into the system. The fact that we have carved out a comparative utopia involving tennis rackets and gaming consoles is a puzzling exception to a long history of strife and it may not last very long. Even if it does, we will remain trapped in the game of life, beset on all sides by viruses and bacteria and other humans who are motivated to seize the niche that we occupy.

Benatar concludes that there is no way out of this game without refusing to play. He has a point. But what if the game could instead be transcended — or even won?

## VII. The Solution

As I see it, the only way to acknowledge Benatar's predicament without accepting its conclusion is to adopt the goal of fundamentally transforming the human experience through technology. Something like: defeat death, conquer pain, overcome biology, explore the universe.

There are a number of ideologies that have similar goals. Benatar describes it as "posthumanist," which works; I have a particular affinity for "[xenofeminist](#)" as a programmer and a former Gender Studies major. But I'm going to use the label "transhumanist," since I think it will be the most familiar to readers, and I like the way it shares a subversive wavelength with "transgender" and "transsexual".

Yes, the lives we live now are painful, short, and cosmically meaningless. But what little knowledge we've been able to glean so far suggests that we have at least the capacity, and perhaps even the duty, to someday solve these problems for ourselves. This, to me, is the essence of transhumanism.

Perhaps unsurprisingly, Benatar anticipates the transhumanist position, and he's pessimistic. To him, transhumanism is "engaged in another kind of secular optimistic theodicy" (89) by promising that a solution to the predicament will eventually be revealed to the faithful, just like most religions do. If Christians justify suffering with a vision of the pearly gates, transhumanists are invested in a similar lie, only with a manufactured heaven full of wireheading whole-brain emulations instead of angels.

There's truth to Benatar's characterization of transhumanists. Transhumanism requires faith: in the idea that current technological trends will continue, that bioengineering is tractable, and that we won't all kill each other before we figure it out. But contrary to traditional religion, belief in transhumanism is grounded in the material evidence of a few thousand years of history.

One of my favorite TV shows is [Tudor Monastery Farm](#), a BBC docuseries where a group of British historians live on a 1500s farm in Sussex and show how tenant farmers lived and worked. *Tudor Monastery Farm* is set about five hundred years ago, which means the time between then and now represents about 10% of recorded history, or 0.25% of human history.

The first impression that you get watching historians cosplay Tudor farmers is that most things in Tudor life were either horrendously uncomfortable or flat-out dangerous. Doing laundry took a full week of work involving poisonous chemicals and hours spent bashing wet clothes with a wooden paddle. Houses were cold and crowded, heated by a wood fire burning in the middle of the room. Even if you were lucky not to lose your hand-sown crop to pests and disease, winters were still long, dark, and hungry. With a few precious exceptions, like salt, you had to know how to make everything you needed by hand from materials that you could grow or gather.

But the second impression you get from the series is that life was getting better for Tudor farmers, in miraculous ways — and fast. Markets were being opened up to allow farmers to buy

and sell goods without permission from the Church or the lords. Books were being printed on a mass scale, creating opportunities for farmers to share knowledge of their craft (and future historians to replicate it). Proto-industrial machines like mills and blast furnaces were producing sophisticated materials with incredible efficiency. The world was gradually becoming more comfortable and more legible.

Benatar would probably point out that in spite of our comfort and our knowledge, hedonic habituation suggests that we're probably not much happier than Tudor farmers were. Fair enough. But in just five hundred years we've developed an extraordinary capacity to *understand* this problem. We have some knowledge of how our psychology works and we have biological theories to explain it. We have experimental methods to test out ways of tweaking our bodies. We've built thinking machines that don't share our psychological problems and we're working on giving them consciousness.

Maybe you think conscious machines and the transcendence of the natural world are science fiction. But if we extrapolate the trendline of the past five hundred years of technology and assume that we don't cause our own extinction in the near future, it would be more surprising if we *didn't* achieve these things in the next five hundred.

## VIII. The Uncertain Future

There's another problem with Benatar's argument that transhumanism can help address: the voluntary extinction of humanity might solve the human predicament, but it wouldn't do anything for the rest of life on earth.

As best we can tell, most sentient beings face a similar predicament as we do. Animals live short and painful lives that have no cosmic purpose. In fact, there's some evidence to suggest that [the animal predicament is even worse than the human predicament](#).

Can humans or their descendants solve the animal predicament? This seems even less clear than the question of whether we can solve our own predicament. Productively intervening on ecosystems will likely prove even harder than intervening on our own biology. But if we were to voluntarily go extinct, the question would be settled. Suffering would conquer all non-human biological life from the day of our extinction to the day our sun finally burns out.

From this perspective, we can see that humanity as a species faces a similar predicament that each individual human does. Now that homo sapiens has been introduced to the world, we've grown attached to it. We've developed scientific and ethical systems for understanding it that can't exist without us. The annihilation of our species would represent the permanent closing of any future opportunity we have to improve this world — the surrender of intelligence to the blind and cruel force of natural selection.

Benatar makes a persuasive case that we are trapped by the human predicament and that we will never solve it. But what if he's wrong? The stakes couldn't possibly be higher. We are the

only hope for this planet — [and possibly the observable universe](#) — to solve the problems of biological life and make a just world for all sentient beings.

Perhaps one day we'll discover that the moral asymmetry between pain and pleasure is wrong, and that we actually do have a duty to stuff the universe full of ecstatic consciousness. Who can be confident either way? This uncertainty alone should make us hesitate. One thing is certain: we won't figure it out without more people to help.

# The Joy Luck Club by Amy Tan

## I. Introduction: The Words You Cannot Say

Have you ever begun to telling a story from your childhood, only to stop mid-sentence?

Maybe everything seemed fine at first. But then you noticed that to reveal the crux of the matter - **why** the story was even funny, **why** the situation was so unexpected, **why** the advice given to you was positively brilliant - you would stumble over an uncomfortable detail. A thing normal and accepted in its place and time, but which - in the here and now - falls into one of a variety of categories:

Impossible. Unacceptable. Just plain weird.

(Even worse are the snippets that slip past such triage attempts. You say a thing, and instantly regret it. It's too late though, because your listener's ears have caught your meaning, and her eyes reflect back to you an inescapable verdict: "What is 'normal' to you is not normal at all.")

For some of us, almost everything that matters about who we are & where we came from is illegible to those we love.

## II. Wounds: The Rifts Between Generations and Cultures

This illegibility of identity - and the consequent impotence to influence - marks all the protagonists of "The Joy Luck Club," a master work by novelist Amy Tan.

The vision that prompted the creation of "The Joy Luck Club" was a vision of desperation:

*"My mother could sense that the women of these families... had unspeakable tragedies they had left behind in China and hopes they couldn't begin to express in their fragile English."*

The story begins and ends with four mothers, and four daughters in and around 1980's-era San Francisco. Each daughter has some piece of her life on the verge of falling

apart. And each mother holds in keeping treasures of memory, wisdom and experience-- but wonders how she will ever pass them on.

In a sort of Chinese Fairy Tale that begins the book, and reflects one life, Tan speaks of a woman who purchased a swan from a market vendor for a foolish sum:

*“This bird, boasted the market vendor, was once a duck that stretched its neck in hopes of becoming a goose, and now look!—it is too beautiful to eat.*

*Then the woman and the swan sailed across an ocean many thousands of li wide, stretching their necks toward America. On her journey she cooed to the swan: “In America I will have a daughter just like me.... Over there nobody will look down on her, because I will make her speak only perfect American English. And over there she will always be too full to swallow any sorrow! She will know my meaning, because I will give her this swan—a creature that became more than what was hoped for.”*

*Now the woman was old. And she had a daughter who grew up speaking only English and swallowing more Coca-Cola than sorrow. For a long time now the woman had wanted to give her daughter the single swan feather and tell her, “This feather may look worthless, but it comes from afar and carries with it all my good intentions.” And she waited, year after year, for the day she could tell her daughter this in perfect American English.”*

But the gifts of language are unevenly distributed among the two generations. It is as though the daughters have access to a set of symbols comparable to ASCII, but the mothers have a set of symbols as rich as Unicode. (In a time and place where ASCII dominates, being adept with those symbols has certain advantages! But if you can't load the extended characters your mom is trying to communicate with, you won't be able to understand her.)

*“She said the two soups were almost the same, **chabudwo**. Or maybe she said **butong**, not the same thing at all. It was one of those Chinese expressions that means the better half of mixed intentions. I can never remember things I didn't understand in the first place.”*

(Chinese words were not using pin yin, but an earlier Romanization.)

Even the names of the mothers can be a barrier: a reader coming from English-language context might find them difficult to track. The mothers are Suyuan Woo - a woman of a long-treasured hope, An-mei Hsu, the thrice-scarred, Lindo Jong, who is like a fierce wind, and Ying-ying St. Clair, who was a tiger, but is now a ghost.

Their daughters have more familiar names: June Woo, Rose Hsu-Jordan, Waverly Jong, Lena St. Clair. Well, mostly more familiar. Waverly was named after the street her family lived on: It's such an odd detail, but you can bet you finally get an explanation of the name - after you've been bothered by it like twenty times!

Foreign language and foreign-sounding names are a barrier, but not the only barrier.

Sometimes, we think that most other people have lives like ours, and live in familiar places. If a Joy Luck Club mother told a story that was resonant with the assumptions of a traditional society, her daughter would probably ask, "Why didn't you just X?" Sometimes the answer was, "There was no X"; other times, it was, "You couldn't 'just X' - what about your family?"

To truly understand, sometimes we need to think the unthinkable.

For at least one daughter, the terror is that there is truth behind the stories.

YingYing St. Clair tells her daughter Lena a story about an ancestor who sentenced a beggar to "die in the worst possible way," and the beggar's ghost returning for vengeance. ("Either that, or he died of influenza one week later.")

Lena's child-imagination, though, enfleshes the scene with particulars, playing it in in her mind again and again. She envisions a beggar condemned to "the death of a thousand cuts." The executioner raises a sharp sword, the beggar perishes gruesomely before the blade even touches his flesh, and the man's ghost returns to haunt both her ancestor's study.

The struggle with terror is not isolated to one family. Rose Hsu grapples with teenage survivor guilt when a family outing to the beach ends in devastating loss. Rose's anguish turned to horror as she was forced to stand as witness first to her mother's denial, and then her particular style of bargaining. Her mother initiated a negotiation first with 'God with a capital G,' (white leatherette Bible in hand) then - in rapid succession - with another, quite different spirit, which she names as "Coiling Dragon."

Thanks to Authorial Sovereignty, we as readers can get a privileged, fly-on-the-wall view that goes back one or two generations. We can see the mothers' early lives in China - and in most cases, the mothers' childhoods, and their relationships with their own families-of-origin. And their own mothers.

We see things that defy explanation.

We see the harrowing experiences that marked and devastated these mothers decades ago. Sometimes they docilely accept the unacceptable so they can survive circumstances that: a reputation assassination that pushes a pious widow into a marriage to a man she despises, as his much-looked-down-upon fourth wife. Sometimes, in despair, they flee from the invading army, traveling towards the hope of safety and casting aside everything that was in their hands - even the one or two things they most needed to save.

But other times, we see the moment where one of these women does something different: she takes her agency into her own hands; she finds a way out of a hopeless conundrum; she plans and she plots and she times her moment just right and seizes the advantage against her adversaries. And she forms this life-shaking solution within the bounds of her own culture: using the tools, the assumptions, and a savvy awareness of the worldview. Though there's certainly room for her own special brand of creativity.

But there's also room for ghosts. A repeated motif in Amy Tan's stories is a functionally-powerless wife (or concubine) threatening to return as a ghost if her would-be tormentors (who are among her own household) go too far - a last desperate weapon of a woman who can't find a way to push against a formidable power differential.

As you read these accounts, it occurs to you that the mothers, as children, believed they were disposable daughters.

### III. Interlude: Using the Category of Concrete-Relational Thinking to Decrypt the Narratives in "The Joy Luck Club"

*"...we are concerned with people who attempt to discover truth in such a way that 'life and reality are seen pictorially in terms of the active emotional relationships present in a concrete situation.' In verbal communication, the concrete relational thinker tends to express, inform, and persuade by referring to symbols, stories, events, objects and so forth, rather than to general propositions and principles. But he is especially prone to rely on nonverbal communication of all types..."*

- "Communicating Christ Cross-Culturally," David J. Hesselgrave

The category of "concrete-relational thinking" as opposed to western propositional-logic-based thinking was eye-opening for discovering meaning within "The Joy Luck Club."

Concrete thinking as opposed to abstract thinking is a concept we are familiar with - but what of the "relational" part? Think of it like this: There is no such thing as a single



parallel line. "Parallel" only has meaning in the context of comparison to another line. It's connected to a focus on the relation of individual elements to each-other: where the people are standing in relation to each-other, or what is the placement of the strokes of a Chinese character.

An example:

In my own life, I, too, have been on the receiving end of the inscrutable wisdom of a Chinese mother. Five years ago, when I was going through treatments for breast cancer, three friends came to my home to offer wisdom, encouragement-- and help cleaning my house!

One of them who'd survived cancer was reminding me about looking after my needs. She ticked off on her fingers, "Physical, Emotional, and Spiritual," when an excited look came into her eyes. "It's like when you are taking a picture," she began, grasping an invisible camera, building a comparison to the three legs of a tripod. Then she quickly concluded, "And that way, you can have the whole family in your picture!" - satisfied that the connection was obvious.

At the time, I flagged it as "something that must make sense within concrete-relational thinking patterns but not within typical Western thinking patterns." Many times since then, I have wondered what my friend could have meant. What was the connection?

Why does the tripod have three legs? "So it will be stable." (as anyone who read that one Encyclopedia Brown mystery knows.)

But that is where we expect a Western thinker's analogy (a metaphor in the middle of a lecture, or a sermon illustration perhaps?) to end. Stability. It's good thing, but is the abstract concept of "stability" a good unto itself? Why not press further and ask, "What is the stability FOR?"

Today, I think for the first time, I finally understood. When a family uses a tripod, it's so that the one taking the picture is not left out of it. And this was her wish for me: she wanted me to be able to be "in the picture" with my family, and remain there.

#### IV. Restored beyond Expectation: Possessors of Agency and of an Uncommon Vision

*"Besides a common language... there will also be, within the same group, many specialized vocabularies... Such 'working languages' are filled with jargon and idioms. People in the same profession typically have a language of the trade. Other people*

*cannot understand the conversations... In every school and even every dormitory room, specialized vocabularies develop. The most widely occurring 'working language' is that which develops between mothers and their children."* -Fei Xiaotong, "From the Soil: The Foundations of Chinese Society"

The unusual thinking styles - not to mention the otherworldly terrors in their life histories does not make for the mothers' easy assimilation into a materialistic American life. The mothers deeply-held desire to be understood is palpable - but it seems that they left their words an ocean away.

However, in the times of testing that adulthood imposes, the daughters had discoveries to make. First a discovery of their own limitations. But second, the discovery that their mothers, whom they have come to see clueless, verbally-clumsy, and incomprehensible are actually strikingly perceptive. It's an awkward dance towards one-another, on a road mostly dark, but occasionally lit by a brilliant flash of illumination.

Sometimes it takes the form of a serious mother-daughter conversation, bracketed by stressful and exasperating interactions.

In response to a shaken Waverly saying, "I don't know what's inside me right now," Lindo tells her, "Then I will tell you." Lindo gives her daughter an origin story of Waverly's life. Who she was, in her mothers' eyes, has everything to do with where her mother comes from and her father comes from.

Then Waverly puts forward a glib comment - the kind of joke people make when they are uncomfortable because they don't understand a thing, or want to test whether they do understand- but carve out a "safe" path they can retreat back down if not. But her words are taken seriously by the mother momentarily, Lindo questions what Waverly means, and then the truth is out.

*And just like that, the fragile connection we were starting to build snapped.*

They can't move forward that way; there is a cost to understanding, and it's full of risk.

*...We sank into silence, a stalemate.*

But with creativity, the mother puts forth another way of thinking about it, and weaves in a joke. She is reaching across to her daughter and she finds her.

*And then her eyes lighted up. "Now listen. You can also say the name of Taiyuan is Bing. Everyone from that city calls it that. Easier for you to say. Bing, it is a nickname."*

*She wrote down the character, and I nodded as if this made everything perfectly clear. "The same as here," she added in English. "You call Apple for New York. Frisco for San Francisco." "Nobody calls San Francisco that!" I said, laughing. "People who call it that don't know any better."*

*"Now you understand my meaning," said my mother triumphantly. I smiled. And really, I did understand finally. Not what she had just said. But what had been true all along. I saw what I had been fighting for: It was for me, a scared child, who had run away a long time ago to what I had imagined was a safer place...*

Mothers and daughters alike are finally seeing things as they are - things they could not perceive before.

Before, Suyuan and Lindo constantly compared June and Waverly - and the old rivalry has never ended.

While Waverly resembles her mother in independence and tenacity, Suyuan's daughter June has inherited her mom's intuitive way of thinking, and her eye for subtleties. She walks into a room and sees things without being told:

*It's her place on the table. Without having anyone tell me, I know her corner on the table was the East. The East is where things begin, my mother once told me, the direction from which the sun rises, where the wind comes from.*

And she's right.

But even more, when June sits down with the three remaining Joy Luck "aunties," she sees the undercurrents in their conversations... she sees the pain written on An-mei's face when a bit of gossip spilled out hits too close to home. She sees Lindo respond to try to "save" the conversation, appearing to talk about one situation, but communicating to An-mei, "You didn't deserve that" about her own bad situation.

## Wordplay, Structure, and, Imagery in "The Joy Luck Club"

On a second read, I began to see the artistry of how "The Joy Luck Club" was put together: it not only bears its own peculiar burden of (riveting and emotional!) truth: it's also sublimely beautiful.

A vase on a wobbly table is suddenly a metaphor for the issue in a daughter's life that is unstable and likely to fail if she doesn't do anything:

*"The only decoration is an odd-looking piece right next to the bed: an end table made out of a slab of unevenly cut marble and thin crisscrosses of black lacquer wood for the legs. My mother puts her handbag on the table and the cylindrical black vase on top starts to wobble. The freesias in the vase quiver.*

*"Careful, it's not too sturdy," I say. The table is a poorly designed piece that Harold made in his student days.*

Any object in a story could tell us something about a situation, if viewed with the right eyes. The structure and the analogies suits the mothers' intuitive way of connecting ideas and concrete-relational way explaining their thoughts. The beauty and structure is so much needed to weather the journey of these womens' life experiences.

Sometimes Tan concisely drops devastating truth into our waiting ears:

*"In two years' time, my scar became pale and shiny and I had no memory of my mother. That is the way it is with a wound. The wound begins to close in on itself, to protect what is hurting so much..."*

Other times, the author litters our path with clues.

She lays out a tapestry replete with repeated words, repeated themes. Sometimes reversed, sometimes translated from one language to the other, sometimes a homonym or a mistaken word that fits oh-too-perfectly. Sometimes you are in one woman's life story; sometimes you find the same theme - but different - in another's. This is skilled wordplay; here is room for intellectual challenge (and emotional challenge, too). It is a joy simply to *track* with the author - to say, "I see what you did there," or then to come back years later and notice - amid the many layers - something you didn't see on the first or second reading.

The overall structure of the book includes a strange or brave choice - Tan divided her book into four sections. "Four" is a very taboo number in Chinese, because it sounds like the verb "die." Not only is the entire book divided into four sections, but each section is divided into four chapters - each one a mother or daughter. It would seem as though this structure is screaming, "Die, die, die, die." While that kind of fits the unrelenting bleakness of the hard parts of these lives portrayed, I don't think Tan is exactly eager to flout tradition. (She didn't number the sections or chapters.) And here's something where I'm not sure if I'm right - but I imagine one of her protagonists - a Chinese mother perhaps - explaining that you can see it in another way. There are two groups of eight sections - and eight is a very *good* number in that system. The daughters complete the mothers' stories, and what was like death without them becomes success.

## V. Conclusion: Catharsis. Cost. Catalysis.

If you've ever had a Chinese friend (or maybe been that Chinese friend) brush aside some unreasonable-sounding parental restriction or awful-sounding phenomenon of their family life by saying "Ah, it's just Asian parents." Maybe some of these are impossible to disentangle with anything short of a novel. (Or maybe they are only impossible to explain to someone unmotivated, and a novel helps many to find the needed motivation.)

When author Flannery O'Connor cast her prophetic gaze upon American readers' moral and intellectual landscape - just like a stereotypical Chinese mother - she saw shortcomings and dangers.

She said:

*"There is something in us, as storytellers and as listeners to stories, that demands the redemptive act, that demands that what falls at least be offered the chance to be restored. The reader of today looks for this motion, and rightly so, but what he has forgotten is the cost of it."*

Amid the confusion, griefs, disasters, and messiness of people's real lives, Amy Tan offers that redemptive act. And where redemption can not be yet seen, she offers hope for redemption. And where even a flicker of hope for redemption can be seen yet, she offers understanding. And she has not forgotten the cost.

# The Millionaire Next Door by Thomas J. Stanley and William D. Danko (and other books)

## You're Probably Wondering Why I Called You Here Today

In 2016, Neal Gabler wrote an [article](#) about how over 47% of Americans [didn't have enough money](#) to meet a \$400 emergency. This didn't surprise Gabler because he was one of the 47%. In the article he recounts his financial history, every now and then stepping out of the narrative to point out a mistake, as if to say, this is where I went wrong. What is fascinating is that almost every time he does this, he misidentifies the mistake. He is a reverse-oracle: he almost always gets it wrong.

At one point, Gabler and his wife own an apartment in a co-op building in Brooklyn. Then they move to the Hamptons and rent a home. A big payday arrives. He uses the payday to put a downpayment on the home he's renting. He has difficulty selling the apartment (the co-op board keeps rejecting buyers), which means he carries two mortgages for years. He concludes that he should have cut the apartment's price sooner.

But was this the real mistake? Could he credibly have known what the right price was? Pricing a complicated real estate asset is hard, even for experts. No, Gabler's real mistake was using the payday as a downpayment. Prior to spending that, Gabler had *years of rent* in his bank account. He was *indestructible*. What can harm you, financially, if you have years of rent in your bank account? You could lose your job and have time to find another, could get sued and defend yourself, would never need to worry about your health insurance lapsing, and if you had trouble selling some real estate like a complicated Brooklyn apartment in a co-op you could just...wait it out. Gabler was invincible, and he chose to *become vulnerable*. And what did he get for it? More debt. He cut off his hand to shackle his wrist.

Gabler, to his credit, admits he is a "financial ignoramus" but he's not a stranger to research; he's written biographies. It made me wonder what research, specifically personal finance research, he'd do for this article, and what he'd found. What are the personal finance resources, and are they any good?

To answer this question I read a few books. Most of them I won't name, because they were terrible. Two of them dedicate a chapter to estate taxes. For estate taxes to kick in you need over \$11 million. \$11 million! Hire an accountant! One author claims to have coined the

term “bad debt”, and argues that land’s limited supply makes it a good investment. Air is in limited supply! Air is free! Horrifyingly, many of these authors professionally talk or teach people about personal finance. The prospective student is in some trouble.

But three of the books I read I took from the r/personalfinance subreddit, and these were, well, they were ok. They are *The Millionaire Next Door*, *Your Money or Your Life*, and *I Will Teach You to be Rich*.

## The Millionaire Next Door

“Spend less money” is good advice, and [like all good advice](#) it is obvious; the hard part is following it. The books have different ways of trying to get you to follow it, and *The Millionaire Next Door*’s way is to show you a bunch of millionaires who don’t spend much.

*Millionaire* is written by Tom Stanley and Will Danko, two PhDs who mainly study rich people. The book mostly focuses on big savers, who have a lot of wealth relative to income (upper 25%), and big spenders, who are the opposite (lower 25%). Sadly the middle half of people are mostly neglected.

The big message of *Millionaire* is that most millionaires don’t seem like millionaires. They don’t consume conspicuously. The book likes the phrase “big hat, no cattle” to describe the opposite: people who earn a lot, and spend a lot, often visibly, but don’t have much money left over. Given this, the big savers behave as you’d imagine. They seek value for their money. They have nice stuff but got it on sale; they live in a modest or nice house and drive a good but generally used car. They don’t buy more than they need. They almost always plan out their spending, which is part of what enables them to get good value for money. They’re pretty boring.

The big spenders are not boring at all. One was a surgeon who spent \$72k a year on cars (mostly Porsches), which is insane. I mean it’s insane now, but this was in 1996. You could buy three new, *loaded* Honda Accords in 1996 every year and still not spend that much. What’s more is that the surgeon was incredibly focused on getting a great price for these cars. He knew the costs for every Porsche dealer in 100 miles. This sort of information takes effort to collect. Naturally, he had very little wealth. The money came in and went right out.

Stupidity can be entertaining, and one doesn’t feel much guilt laughing at a surgeon making \$700k a year. But it struck me that there can be a public cost to this ignorance. Taxes are bad, but if you have to tax, better to mostly tax the rich. The rich have more money, and taxes are less likely to impact their day-to-day life. But if they always spend too much, it will

impact their day-to-day life. A surgeon doing very effective research on Porsche dealerships in order to support his dumb spending habit is amusing. But what if he's doing very effective political activism to support it? That's not so funny. I'm not saying rich people's taxes should go up. Maybe they should, but tax is a hard problem. But probably "I don't want to give the public more money because I'm bad with money" is not what that debate should turn on.

*Millionaire* excels in storytelling but falls short as a guide. The central argument, "spend less money", is delivered fairly convincingly, but the book doesn't offer much guidance as to *how* you should do this, apart from basic advice like planning your spending, having your spouse on board, and avoiding environments that encourage consumption. It certainly won't help you get out of debt.

Still, *Millionaire* only claims that it will tell you *about* the wealthy, not how to join them. Does it? Stanley and Danko get most of their data from interviews and surveys; they mail surveys (\$1 enclosed) to rich neighborhoods. For the interviews they seem to offer \$100 to \$200 for an hour or so. I doubt this is enough money. The interviews are live and require the subjects to travel, which puts the pay rate between \$50 to \$100 an hour. Even in 1996 that is low for a person with a million in the bank. A dollar seems similarly puny to read a letter, fill out a survey and mail it. It suggests a sample bias toward those who don't understand the value of their time, or whose time is much less valuable than their assets might suggest. Some details of the big savers are consistent with such a bias.

Many of the big savers behave as if saving is a compulsion. One woman continues to cut coupons after her husband gives her \$8 million. I may not know the dollars-per-hour of clipping coupons, but I do know that a 5% annual return on \$8 million works out to \$200 per hour. Now, my economics professors might suggest that this woman just so enjoys clipping coupons that for her clipping coupons is really consumption rather than production, but actually they would never say that because my professors were *good* professors and the idea of clipping coupons is nonsensical, it's *insane*, read a *book*, go on a *hike*, watch *Queen's Gambit*, we live in a golden age and you have eight million *dollars*, stop clipping coupons and live your *life*!

Big savers also invest strangely. They go through a lot of trouble to avoid taxes by, say, investing in tax-free bonds instead of stock market index funds. This is a terrible move, because even if you somehow found a tax-free bond yielding 5%, you'd still be way behind an after-tax yield of 8%. Five and eight might seem like similar numbers, but 1.05 to the thirtieth is about four, whereas 1.08 to the thirtieth is about *ten*. That's \$400k in retirement versus \$1 million. It's a big difference!



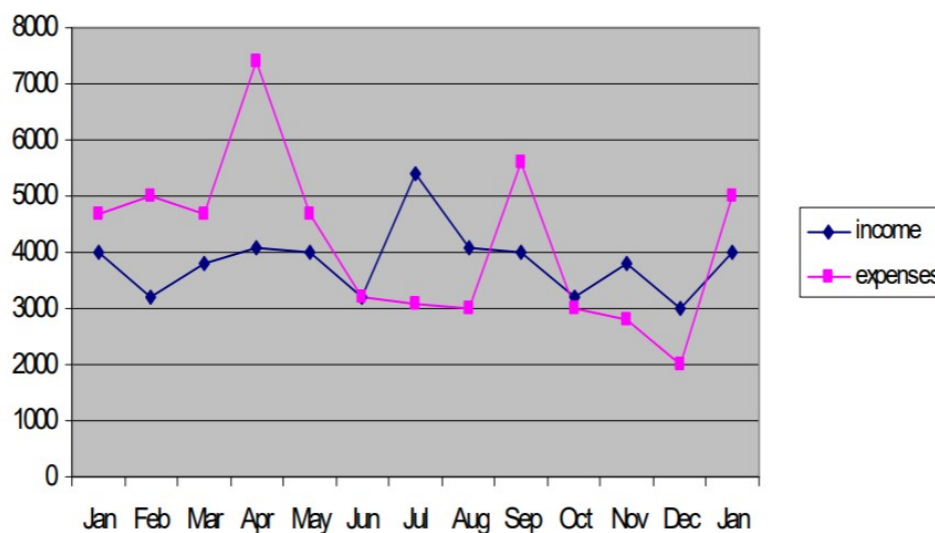
Two chapters are on rich parents giving their adult children financial support. Usually the kids blow it, and the book concludes that doing this is always terrible. I'm skeptical. Surely if someone is properly equipped to use money, then more money will be helpful. Properly equipping someone might indeed be very hard, but it isn't impossible; I've seen it happen. It happened to me. But for someone who doesn't have the information they need to impart, it wouldn't just be hard. It would be impossible.

## Your Money or Your Life

*Life* is the oldest of the books, so old that it has outlived one of its authors. It seems to be the origin of the FIRE movement (Financial Independence, Retire Early), a community/movement of people who try to retire ridiculously early (like, in their 30s) by saving huge portions of their paycheck and living very frugally.

According to *Life*, you've been thinking about money all wrong: it isn't really *money*, but *time*. You trade time for money when you work, and your time is limited, so you'd better be damn sure you're getting a good trade. *Life* is the most specifically prescriptive of the books, and clearly details how to reduce your spending, increase your income, and eventually never have to work again.

The method has two parts. First, the book suggests you track and categorize, every month, every dollar you spend. Then, ask yourself: do I feel good about that spending? Finally, adjust your spending until you are only spending on things you feel good about. The second part is to make a chart that records your income and your expenses each month. Like so:



This a “wall chart,” because you’re supposed to put it on your wall, ideally one you look at daily, so that it can help you stay motivated to keep hold of your money. Eventually, you’d add a third line to the chart for your investment income. Eventually, through the magic of compound interest, the investment line grows enough to cross the expenses line, and then, [you’re free!](#)

You don’t have to work anymore. You exist from a position of [fuck you](#).

Stylistically, *Life* excels at putting things in perspective. The book asks you to figure out what your real hourly rate is, by factoring in work expenses both in time (commutes, time decompressing after a long day, etc.) and money (office clothes, commute again, and so on).

This often results in a much lower hourly rate than your salary would suggest. When you go to track your monthly expenses, *Life* encourages thinking of those expenses not as dollars, but as the time you traded for them. This is the right way to look at it. The stumbles come from *Life*’s various tangents, which are numerous, and make reading it a real slog. Many of the claims in these tangents are by this point so well-refuted that they have become [punchlines](#). Some are just lazy. At one point, the book talks up sustainable investment funds, and cites as evidence an article from a website owned by a bank that *sells sustainable investment funds*.

When not distracted by tangents, *Life* gives the best advice of all the books. It is the only book to emphasize an emergency fund, which is critical. It misses in two areas. The big one is that the book never tells you to hire anyone; in fact it recommends you try to do as much as you can yourself. This is really disappointing. *Life* recognizes that you can trade time for money, but doesn’t understand that by hiring people, you can trade money for time. Hiring someone to clean and do laundry once a week is not expensive; and when you are rich, it is not expensive to have them come several times a week, or to also cook for you. It may be less than what you already spend on delivery and restaurants. If chores are a source of tension between you and your spouse or significant other, then my god, for a little bit of money you can have peace in your house. If you have a job you enjoy and chores you do not, then you could do more of the job you enjoy in exchange for never doing chores you hate *ever again*.

In investing, *Life* is better, but still not perfect. The authors admit that you should put your money in index funds, but only grudgingly, and the risks are overemphasized. Real estate gets a lot of space, even though it’s generally a bad investment. A lot of the investing section involves the authors reminiscing about how government bonds used to be better in the good old days, while conveniently forgetting that there was enormous inflation in those not-actually-good-at-all old days which canceled out the yield on the bonds, and anyway the indexes still did much better than the bonds even then.

*Life* never really explains why index funds or stocks in general are such great investments. I feel that explaining *why* stocks are good is very important, because without that understanding, it is very difficult to own stocks without panicking every time the market goes down. The explanation is simple: when you own an index fund, what you really own is a piece of the output of the *best people in America*: the smartest, the hardest-working, the most capable, the best trained and the best equipped. There are of course some great people that you aren't catching, but that's fine; there's plenty to go around. If you own an index, Bill Gates, Jeff Bezos, Elon Musk, Warren Buffet and millions of amazing people you've never heard of are all striving to put money in your pocket. They put it there by getting things to people that need or want them, by enriching lives and by changing the world. Again, there are some exceptions, but they are small and you can ignore them. Stocks are not something magical or mysterious; they are not numbers that go up and down, or squiggly lines on charts. They are the best pieces of the best economy on Earth. *That* is what pays the dividends and what makes the prices go up. It is powerful. It is reliable. If you can understand this, if you really get it, if you can feel it in your bones, then you are already rich: it is only a matter of pressure and time.

## I Will Teach You To Be Rich

One of the downsides of getting older is that you encounter things you liked when you were younger, and you realize what incredibly poor taste you had. That was my experience with this last book. *Rich* is based on a [blog of the same name](#) that I used to read. Ramit Sethi is the writer, and yes, that is his enormous picture on the front of the blog, screaming the unmistakable impression: "I'm a giant douche". His book does little to contradict this. A disturbing amount of text is just him bullying customer service. At one point he challenges people to save \$1,000 in one month, and when some of his audience point out that they do not *make* \$1,000 a month because they are a student or housewife or just really poor, he claims that saving \$1,000 is an "aspirational goal" (he did not say that before) and that they should try to save \$500 or something and then harangues them for needing to be told this. He encourages you to "dominate your clueless friends". He talks about how he asked his wife, to whom he is still married, for a prenup, in a book for strangers. I am reading *his own account of events* and he comes off as a jerk. I didn't understand "punchable face" before, but now, I know.

Despite this he does say some good things, often more specifically than the other two books. Also, in defense of my younger self, the front of his blog didn't look that way when I first encountered it, and I used to be really stupid.

Instead of a really intensive expense and investment tracking system, Ramit suggests an "85% solution" that takes much less time but still gets most of the way there. It's probably the second-best book (*Life* still does a better job) in terms of the quality of its advice, and also the best-written. Ramit is a blogger and it shows. He is concise, snappy and clear. Instead of giving long-winded summaries of individual cases as the other books do, Ramit quotes them directly, briefly, and cannily. I was moved by the story of the young man who'd gotten serious with his girlfriend, discovered she had saved far more than him despite making much less, and resolved to change. Suitably, the best section of this book by a person who professionally talks about money is his advice on how to talk about money, which I highly recommend. In a way it is a shame that Ramit is such a capable writer, because you get a sense, reading him, of what a really great book this could have been, if only he'd been able to restrain himself.

Unlike the previous books, Ramit correctly observes that real estate is often a bad investment, and that there's nothing wrong with renting, especially if you're not sure you can commit to a long-term stay. He acknowledges that some spending is good, though he misses the benefits of hiring people. The section on debt and getting out of it is excellent. He points out that you can often reduce your credit card debt simply by asking your creditors for a better rate, especially if you have leverage, like the threat of transferring the balance to another card. He details an "envelope" system for managing spending, where you put what you intend to spend in several categories in some envelopes (which can be virtual), and only spend from those envelopes (you can move money between envelopes, but you can't add any more) which is simple, easy and effective. He focuses heavily on automating as much of your savings as you can, because that makes it more likely that you will actually save. Part of this automation is target date funds, which really do fulfill his "85% solution" ideal. He tells you you should buy index funds, but unlike *Life*, he's enthusiastic about it, rather than grudging. While all the books encourage you to get a raise, Ramit is the only one that offers concrete advice on how to negotiate a higher salary, though I'm a little concerned that parts of his approach are too aggressive.

Ramit doesn't seem to have much of a sense of proportion. *Rich* spends a ton of time on the benefits of credit card spending, even though all of those benefits combined are less than 3% of part of your spending, which is peanuts. He spends another chapter on banks, and a lot of time telling you things about 401(k)s that are mostly irrelevant. *Life*, by contrast, just tells you

to use cards if you can keep yourself from borrowing on them and to get a bank account with no fees.

Some things Ramit just gets wrong. He misunderstands 401(k)s and asset allocation, and mistakenly suggests that there's no difference between paying off student loans faster and putting the money into investments. In his zeal to encourage index funds, he goes too far. To Ramit, it is not simply *hard* to pick stocks yourself, but *impossible*. What's really bizarre about this is that he relates a story of how he tried to pick stocks as a kid and was *wildly successful* because he bought Amazon early. To be clear: Ramit does it, and concludes that it cannot be done. And of course it can be done, it is possible to pick stocks successfully. Warren Buffet has done it. Peter Lynch has done it. I've done it. At one point, Ramit ridicules people who think a stock is "going to the moon", implying it is foolish to believe such things can be predicted. One of the companies he ridicules people for believing in is Apple. Another is Tesla. Both went to the moon.

## So What Should I Read?

None of them. The books try, and some are ok! But none are good. What I *would* recommend is actually the r/personalfinance [subreddit](#). They have a wiki, and that wiki is really good. The best of the books I read basically give you some of the advice on that wiki, but badly and longer. The books I didn't name miss by a mile.

In a lot of ways this is discouraging, especially because people who do not think to go to a social media website are likely to get caught in a quicksand bog of terrible advice. Which is strange. Other expert fields do not behave in this way. It would be hard to imagine a physics book, written by a physicist, that gets big parts of the field wrong. Lawyers do a good job describing law. What is so unusual about money that money experts do not understand it?

One possibility is that accuracy is not actually the right way to judge a personal finance book. Often the critical thing about advice is that you follow it, not that it's perfectly correct. It might be that the examples in *Millionaire* are just what a person earning a low six figure salary needs to read in order to get their spending under control, and the fact that they'll get tax free bonds instead of index funds is small compared to the effect of that spending. It might be that the droning examples and weird tangents of *Life* get people to *do the thing*. Maybe Ramit does know that you can pick stocks successfully, but he's so worried that his audience will get in over their heads if he admits it that he pretends it can't be done.

Another possibility is that an understanding of money is just really, really rare. That would certainly explain why the purported experts selling books keep getting things wrong. It would explain the bizarre habits of the big savers detailed in *Millionaire*. It would explain why a man bright enough to make a living writing biographies could be unable to meet a \$400 emergency, and why there could be 140 million people just like him.

Both could be true, and I suspect both are, at least a little. But there is a third possibility, which I think is bigger: that the emperor is naked. In the story of the Emperor's New Clothes, people do not want to admit that they cannot see the emperor's clothes, because doing so might mark them as stupid and bad at their jobs. It is a story of how nonsense can persist even in very visible circumstances if people feel they will be threatened by admitting their ignorance. Consider, further, that if the nonsense *does* persist, people who are willing to admit that they don't know are faced with a public square full of noise that they do not know how to distinguish from signal. It would be like having herd immunity to the truth.

In many of the books, especially the bad ones that I didn't name, signs of this abound. Many of the bad authors spend a startling amount of time (often at the very start) warning their readers not to listen to other sources. Ramit condemns not only strategies that have worked for others, but also for himself. *Millionaire* does not acknowledge that the tax-focused strategies of its sample might be wrong. *Life* refuses to consider that you can trade money for time even after hammering, over and over, that money is time. It isn't just the books. Neal Gabler's article does not profile a man seeking to learn, but one looking to blame. Immediately after admitting he spent too much, he says that it would have been fine if his income had kept increasing, except of course he probably would have just kept spending. I personally have had the experience of advising someone that they needed to have an emergency fund, only to have them respond angrily that they already knew that, which of course they did not. I know of so many people who will defend having made almost no return on their money by saying, "I haven't lost any." A friend told me about a fellow she knew who hated one kind of yogurt and loved another kind, but bought the kind he hated because it was fifteen cents cheaper. It seems like every few months there is a story of someone, often a celebrity or athlete, or sometimes a professional who should *absolutely know better* who has done something [breathtakingly stupid](#), something one meeting with a good accountant would have prevented completely. These do not seem like something logical; they seem like something emotional. Something animal.

There's a lot about this shortfall in the financial literature and its possible implications that is discouraging. But in other ways, it is encouraging. It's a counterexample of the usual narrative where social media destroys everything it touches and makes life worse.

[r/personalfinance](#) isn't just really good, it also has tens of millions of subscribers, probably much more than any of the books have sold copies. It's also a great example of how easy it is to exalt books (and older media in general) when in reality most books are as flawed and ordinary as their authors. It offers hope that things *are* getting better, that by working together we really *do* make things that transcend, rather than embody, the pettiness and self-regard that dooms so many human endeavors. Indeed, it is like index funds themselves. Here the wisdom of crowds and of markets, markets of upvotes and posts rather than dollars, but markets nonetheless, doing much better than individuals acting alone. It is a story of how apart, we are limited by our weaknesses, but when working together effectively, we are limited only by our strengths. The effect is real: the number who *can* meet the \$400 emergency is going up over time. In 2019 it was no longer forty-seven percent, but thirty-nine. That is twenty-six million people who were in trouble, and are now ok. They are now ok.

And they are getting better.

# The Ministry for the Future by Kim Stanley Robinson

Your concentration is one part per billion. Actually - one part you, per eight billion parts everyone else. Some poisons are effective at around that concentration. For instance, botulinum neurotoxin, botox, is lethal for a 75-kg human at ninety parts per billion. Botox's lethality stems from its specificity, biocompatibility with the animal nervous system, binding to a nerve cell and preventing it from communicating. It is a paralyzing agent. Abstracting the components to focus on the mechanism, if you could paralyze the carbon-chugging global system in an analogous way, then maybe you could stop climate change. That is the idea behind Kim Stanley Robinson's new novel, *The Ministry for the Future*.

Kim Stanley Robinson is one of the greatest sci-fi writers alive today. He is most famous for *The Mars Trilogy*, where a group of a hundred or so earthlings battle against time to terraform Mars. In *The Ministry for the Future* he swivels his telescope to another planet, Earth in the very near future. Specifically, to summer 2025, also known as four years from now. Things have not changed much, geopolitically speaking. Gas is still as cheap as Diet Coke, carbon emissions continue to climb, and no country is on track to meet their 2015 Paris Agreement targets. Climatologically speaking, the Sun has continued to irradiate the earth, exciting particles in the atmosphere, and changing continental pressure systems. A belt across India, Pakistan, and the North China Plain reaches a temperature no longer compatible with living -- a 95 degree Fahrenheit wet bulb temperature in Northern India -- killing more than twenty million people. Frank May, a young American working in an NGO, leaves his office after the AC sputters out after the generator gets stolen after the electric grid fails. He goes to the river. He jumps in. Suspended in the Ganges with hundreds of dead and dying Indian men, women, and children, he gets poached like a soft boiled egg. For whatever reason - luck, hydration, or body mass privilege - Frank survives, a lone white guy in a river of corpses.

Harrowed by guilt and heat-triggered PTSD, Frank wanders an ambivalent planet, from Glasgow, to Antarctica, to Zurich, facing the challenges encountered by every messiah and conspiracy theorist. He needs people to believe, to feel it in their bones as he does, that not enough is happening, that survival is on the line.

What should he do next? He is just one in eight billion.

## Geoengineering

One way to have an outsized individual impact on the climate is to assume the political changes necessary to address the root cause -- carbon emissions -- won't take place or won't take place fast enough, and take the reins yourself.



First, a cute story. I learned from a friend involved that the US military must a) spend 2 million dollars, and b) submit to a two-year EPA permitting process to remove a tiny grove of trees on Midway Island in the Pacific. They are doing this so they can retire the woman whom they currently employ as a human scarecrow, to keep birds out of the grove, to keep the nearby runway clear of bird droppings. Given this inertia on environmental projects, it may surprise you that atmospheric geoengineering is - for all legal, practical, and social purposes - totally [unregulated](#). That is, until you remember the absence of such constraints on what can be shot into the atmosphere is how we ended up in this hothouse in the first place.

If you don't believe me, then you can just ask Russ George, the billionaire who dumped 100 tons of iron ore off the coast of British Columbia, to 'fertilize the ocean' and bring the salmon run back to the Haida Nation. He faced the Canadian equivalent of a strongly worded [letter](#).

Atmospheric geoengineering - euphemistically phrased 'sunlight management' - ranges from thinning wooly cirrus clouds, to micro-bubbling the surface of the sea, to changing the color of roofs, to '[sending a giant mirror into orbit](#).' But the most comprehensive sunlight management project is dissipating aerosols of calcium carbonate and sulfates into the upper layers of the stratosphere. These white and yellow aerosol particles provide a thick vapor with a reflective enough albedo (color) to repel sunlight back into space.

This process mimics what occurs after a volcanic eruption, and could, in the limit, lower the global average temperature about one degree Celsius. Incidentally, these stinky aerosol particles will scatter sunlight of a greater variety of wavelengths, not just the short wavelengths, and is catching negative PR for changing the sky from blue to white, everywhere. The ocean will remain blue, because it is not [strictly](#) a reflection of the sky. That is, unless we decide to dye it yellow to reflect more sunlight.

Side effects include guaranteed acid rain, potential droughts, and a potential apocalypse thanks to unknown unknowns. As of late February 2021, a team at Harvard has applied to launch the first tests of sunlight management above lapland in Sweden.

The intervention would be relatively cheap - between [two](#) and ten billion USD - and additionally requires one to possess a fleet of jets to ferry the aerosols up 10 km above the surface. Other than those constraints, it works as a unilateral temporary solution to a limited amount of warming. Because it addresses none of the underlying issues causing the warming, it would need to be continued, maybe forever.

But what else can be done? Can one person change the political system?

# Terrorism

Popular climate books and articles tend to cluster along three modes when it comes to messaging strategy. First, there are the doomers, millenarian environmentalists who don't know much about central banking or cement manufacture, but who heavily imply that human civilization in its current form will and should die out, probably. David Wallace-Wells is a big fan (no turbine pun intended) of this sentiment: "We are only just entering our brave new world, one that collapses below us as soon as we set foot on it." Next there are blamers, who appear to think that the primary issue with climate change is that it was caused by conspiratorial fossil fuel companies and rich countries. (Even before reviewing the evidence, I have no doubt that this is true).

Synthesizing these first two messaging strategies used in the current media landscape, Kim Stanley Robinson supposes a future characterized by ecoterrorism as a method of personal expression. In *The Ministry for the Future*, a mythical day, Crash Day, happens where a group of Indian terrorists brings down every plane in the sky with a pack of drones, making flying on planes a lot less palatable. (It should be noted here that Kim has a penchant for steampunk technology, and that dirigibles and blimps figure prominently, using every sense of the word, in the plot). Shipping tankers are also sunk by submarines. Capitalism doesn't fall but it certainly sputters as the heads of fossil fuel companies are systematically assassinated.

To critique this section of the book, one is hard-pressed to find predictions on the efficacy of these types of actions. That is, unless one is familiar with today's most pernicious ecoterrorists: Catholic nuns. The most applicable cases come from the apostolates of Catholic nuns against BlackRock investment, nuclear power plants, and oil pipelines. In 2015, the Adorers of the Blood of Christ successfully blocked a pipeline across their land. Less successfully in 2014 an 82 year old nun broke into the Fort Knox of nuclear weapons material near Oak Ridge National lab to attempt to smear blood on the warheads. Exactly nothing happened politically, she and her co-conspirators got five years prison time, likely 100% of the rest of their lives, for their troubles.

The people alive today, particularly the ones concerned about climate change, are the least religious and least violent group of humans since people bothered to stop disemboweling one another long enough to start keeping [records](#). It seems optimistic (pessimistic?) to assume that we could just turn around and start shooting innocents out of the sky en-masse. Some other option has to be taken - one that is better suited to our talents.

Additionally, it's not clear that we could blast, strike, or otherwise terrorize our way out of the most pressing problems, the ones contributing the most emissions. Consulting Bill Gate's handy [cheatsheet](#) for climate emissions, plane travel, for instance, is not one of the heavy hitters. All of transportation, humans and goods, via land, sea, or plane, accounts for just 16% of global emissions.

This brings us to by far my favorite section of *The Ministry for the Future*, where eco-

libertardians such as myself are forced into a crash course on central banking, exotic financial instruments, and, worst of all, distributed ledgers.

## Financial Engineering

Mary, the hot, 60 year old, Irish authorial surrogate is kidnapped by Frank. Her position in the fictional, bureaucratic, UN-associated Ministry for the Future means that she has to listen. She needs to do more, he says. She believes him because he reminds her of her dead husband. She authorizes a black wing of the ministry for the future. She releases a fleet of climate lawsuits on behalf of the unborn. However, most impactfully, she consults an AI specialist named Jaenis Athena who introduces her to banking and ... the blockchain.

Zooming back to our world, blockchain technologies, bitcoin in particular, are the latest persona non grata in the climate change battle, at least on Twitter. Bitcoin mining is so energy intensive that it has erased the gains of 30 years of solar power. Bitcoin mining today is responsible for 1% of global emissions. Let it be known that this is still dwarfed by the emissions of farting and belching cattle - 4%. Given this, I was surprised at the outsized role a fictional blockchain technology, Carboncoin played in the Ministry for the Future's plan for saving the planet. It is actually not fictional, and it does exist. Carboncoin is a currency that you can 'mine' by reducing carbon emissions. Plant a bunch of trees - 1 Carboncoin. Replace a coal plant with a dam - 12 Carboncoins. (For this ratio, you would have had to have planted A LOT of trees) The magic of the currency in the Ministry for the Future is that Mary, using her feminine wiles, convinces all eight central banks to back the currency. This means that they certify the Carboncoins as a legal tender, thus guaranteeing that at any point you can take it to a bank and exchange it for another currency like gold or something. It is carbon quantitative easing. A carrot to the stick of a carbon tax.

The reason you need the blockchain for this is that people lie and you cannot trust them. Carbon offset markets, where people trade amounts of carbon sequestered, have existed for a long time, since the Kyoto Protocol. However, they have a significant free-rider problem. Under the Kyoto Protocol something like 0.6 billion metric tons of carbon offsets were misallocated, meaning they were given to people who did not actually prevent carbon from getting into the atmosphere. Blockchain uses a secure consensus mechanism to synchronize bookkeeping among parties that do not trust each other, so bad actors cannot 'cook the books'.

Arguments in favor of carbon quantitative easing accounted for on a distributed ledger? Well, first the magic of money itself is that it is 'a social construct.' Fiat currency is widely accepted because of the state monopoly on violence, and before fiat currency, gold was accepted because it was soft and pretty. Central banks have accumulated US\$20 trillion in total assets, indicating that they have the reserves to purchase carbon offsets from petrostates like Saudi Arabia. Blockchain based solutions are already being used to create an ethical diamond market, or rather to price diamonds' origin into the market. At each point in the supply chain, mining,

refining, export, data must be verified on the ledger. ([source](#))

The big argument against Carboncoin is that central banks simply won't do it because it involves taking on significant amounts of risk politically and financially. Some might say that doing nothing to combat climate change is also taking a risk. This solution is out of reach for most people trying to maximally impact the climate system, but, going back to the botox, it provides insight into what is analogous to a nervous system in our world -- the market.

Ultimately, the most important features of The Ministry for the Future's imagined planetary decarbonization were not included in the book. They were the unnamed innovations, construction projects, and ventures unleashed and funded by Carboncoin. The book provides an example of a policy level decision sufficient to enable coordinated efficacy at the local level, but perhaps it is not necessary.

## Something Boring: Nuclear Plants, Green Concrete, and Wheat Semi-Dwarves

Bill Gates has recently refocused the oculus of his intellect from global development to the climate and has released a plan. The book, 'How to Avoid a Climate Disaster' is deeply motivational, particularly because it avoids doomsday scenarios, and because it reveals that Bill Gates is not too old for fart jokes. The tone of the book is kind of like if climate scientists were trapped in a case interview by a consulting firm. The upshot is that an individual or organization can have an enormous impact on our adaptation to climate change, particularly if you diddle some wheat or corn genomes or consider the particulars of cement.

Also unique to his climate proposal is a modicum of concern for the world's poor and a belief that they deserve all the benefits of growth. Electricity, cheap power, is central to increasing productivity, which is key/is economic development. He reiterates throughout that any solution developed must be cheap enough for middle income countries to afford.

To reach his goal of reducing the amount of carbon we release into the atmosphere every year from 51 billion tons now to 0 by 2050, we need to address the top five sources of emission. They are manufacturing things (31%), the electrical grid (27%), growing food and food for our food (19%), transportation (16%), and HVAC (7%). Greening the grid is the single highest impact item, because with cheap, clean electricity, we can cut the emissions of all other categories, provided they are electrified (itself is no small task).

Given the critical place it holds in the solution to climate change, it is no surprise that the electrical grid is the most controversial section of How to Avoid a Climate Disaster. We already have a greener electric grid than I expected. 50% of the UK's electricity is produced by water, wind, or solar. On a good day, 80% of California's is. The problem is a stupid boring problem, and it is that the sun does not always shine and the wind does not always blow, and that we do

not have city-sized batteries just lying around that we can fill to keep for a cloudy day. Battery technology, according to Bill, will probably not improve that much on the Lithium Ion battery that is in the scooters and Teslas. It seems to me like you could just roll a large boulder up a hill and call that your battery, but in these engineering calculations, the details of efficiency, not 'in the limit' behavior is what counts. Lacking storage variable energy sources like wind and solar are buffered by a run-in a CO2-emitting standby mode called spinning reserve. Spinning reserve could be a coal or natural gas plant, or, as suggested by Bill - a nuclear plant.

There was a time, the 1990s, when the FDA approved an electronic brain implant with a 5 year lifetime for Parkinson's patients that was powered by a tiny nuclear power source. It was very safe, it was only discontinued after some unfortunate incidents with corpse handling (Cremation! What a disaster). Public opinion has obviously turned on nuclear power, and the big counterpoint to How to Avoid a Climate Disaster is Professor Mark Jacobson, who in "100% Clean, Renewable Energy and Storage for Everything" (2020) calculates that every country on Earth could afford to get to 80% renewable energy by 2030, that the cost of solar energy has fallen so much that we are 100 years ahead of price projections, and that battery technology prices have fallen along a similar curve. Even very cheap batteries will not meet the needs of large cities. New York City would require a 12 Gigawatt mega battery, and I will need to read the rest of Jacobson's very detailed textbook to get a better picture of his proposal for spinning reserve for a city.

It also seems like maybe they could just turn the power off for non-essential uses during peak demand. Just saying.

Even if we green the grid, there are very large problems to be solved when considering the manufacturing sector. Concrete, the best building material, requires cement, which in turn requires calcium from limestone. Producing a ton of cement produces a ton of carbon dioxide, not from any energy-consuming heating or cooling process but rather from its synthesis. Steel similarly has a very difficult to fix synthesis. Plastic might soon become a net carbon sink.

Climate change was mostly created in just one generation. Half of all emissions have occurred since the first episode of Seinfeld aired. In setting out to write this review, I didn't expect to get a clear answer on what one person should be doing to combat climate change, but I did. In the Ministry for the Future, Frank should have stayed in Denver and worked on the Bessemer process. Or maybe he should have gone to Rhode Island and built more offshore wind farms. He might have also gone to Florida to save the oranges from greening disease (say goodbye to orange juice, friends). In the future, all jobs will likely become implicated in solving climate change, whether you like it or not.

# The No Breakfast Fallacy: Why the Club of Rome was wrong about us running out of resources by Tim Worstall

Having decided that I wanted to review something with a broadly libertarian worldview, it struck me almost immediately that the obvious choice would be Deirdre McCloskey's *Bourgeois Trilogy*, which is monumental (at 2000+ pages), magisterial, and erudite.

Due to the constraints of space, time and competence, however, I chose Tim Worstall's little book *The No Breakfast Fallacy: Why the Club of Rome was wrong about us running out of resources*.

Published in 2015 and comprising just 54 pages of text plus two appendices, it might be better thought of as a pamphlet or a monograph. It was originally sold as a paperback but is now only available online (free), courtesy of the Adam Smith Institute where Worstall is a Senior Fellow.

## Keeping the Focus Narrow

Conversations about mining and resource extraction invariably include discussions of social cost, pollution, and environmental degradation. They also turn to the ruthless behaviour of avaricious oligopolists which is endlessly debated and criticised, all the while threatening to descend into the depths of a culture war with painful predictability.

*The No Breakfast Fallacy* doesn't do any of those things. Tim Worstall in fact makes a particular point of steering away from discussing anything vaguely environmental at all. To give you an idea of what he wants to concern himself with (and the concerns that he admits exist but that he's going to ignore) I'll let him introduce his book himself :-

*"This book is not trying to solve nor even discuss all environmental problems that exist. It's not even trying to discuss all that might occur over the use or not of all minerals and or metals. It doesn't address the possible pollution of areas by mineral processing, it doesn't address nor even attempt to the possible problems of actual use: say, algal blooms from the over use of pesticides.*

*It attempts to discuss one thing and one thing only. Are we likely to run out of any of the minerals or metals that we like to use in anything of a timescale that should be of concern to us today?"*

That's fairly clear.

But is it reasonable? I think so – I have no problem with writers focusing on isolated specific issues, particularly when that's their area of expertise. So, readers with more environmentally based concerns will have to wait for a different book, and a different review. Right here it's all solely about the running out of the stuff of the earth.

## The Fallaciousness of 'No Breakfast'?

The main purpose of Worstall's book is to expose an extremely widespread misunderstanding of the way a single word is used - the word *reserves*, particularly when referring to finite resources such as metals and minerals.

I'm a firm believer that the immortal god of language is usage, and am therefore reluctant to castigate anyone for using words in whichever way they choose. This is true especially when there are hearers (or readers) of the words who understand them in the exact same way they were meant by the speaker (or writer). However, if one group of people - mining engineers, geologists, and commodity traders for instance - use a word (say, the word *reserves*) in one way, and another group of people – everyone else – uses the same word to mean something completely different, then I'll agree with Worstall that this will likely cause considerable confusion.

The 'everyone else' conception of mineral reserves is that they are roughly equivalent to 'how much there is left'. Which is unsurprising given that the word *reserves* in other contexts is used to mean exactly that kind of thing. Confusion arises, though, because the mining engineers and friends use mineral reserves to mean something very different. Utterly and profoundly different! Which is the reason for Worstall's lack of breakfast. Or at least for his coining of the 'No Breakfast' fallacy, to show how almost all of us misunderstand what the mining people mean when they talk about mineral reserves.

Worstall likens mineral reserves to the contents of the world's fridges, in the sense that they are things we have spent money on, and prepared for imminent use. And the error is to think that once we have consumed our mineral reserves, industrial civilisation will collapse and billions will die because our reserves are 'all there is', or 'all there is left'. His contention is that as mineral reserves are much more like the bacon in our fridge, they can be replenished over and over again, so all talk about 'running out' or 'exhaustion' is entirely bogus, just as it would be to say you're going to starve because you've eaten all the bacon in your fridge.

## Renewable Versus Finite

My first minor problem with Worstall's book is that he uses a renewable resource – bacon – as an analogy for finite ones – minerals. And I suppose I bring this up on behalf of my environmentally minded friends - it isn't a huge problem for me personally because, by and large, I agree with almost all of his contentions, especially his main thesis.

Yes, most people misunderstand what the US Geological survey is saying when it claims that the reserves of a particular mineral or metal are a certain number of tons. And indeed this misunderstanding is categorical and profound. And for sure the amount of extra resources that can be turned into reserves is mind-bogglingly huge. And of course the ensuing confusion causes many people to say many things that are utterly nonsensical.

But most people I know who are concerned with resources 'running out' are unimpressed by the argument that there's a lot more of the stuff lying around that can be turned into more reserves. The response to that argument is invariably "Well, they've got to run out eventually, right?"

This strikes me as bound together with our sense of the resources (and their reserves) being *finite*. And there's also an implied contrast with other things that are described as *renewable*. Worstall isn't concerned about finiteness (or renewability, for that matter). His argument is (without wishing to spoil the plot) that there's simply way too much of everything, to worry about stuff 'running out'. An approach which I shall call the "Vastness Gambit".

But I think it's at least worth digging down a little into what we generally mean when we use the words 'renewable' and 'finite'. Like many words we use them partially for signalling and also in emotive ways. Renewables = sustainable = good! Finite resources = unsustainable = bad! But the salient features of such things may entirely pass us by.

Living in a rural environment as I do, I see outside the window of my study some chickens, some horse manure, and some coppiced chestnut trees. The salient features of these things being what? That they can be replaced with more of the same? That they are 'renew-able'? Perhaps. But I'll offer you three similar items from the renewables basket – trees on Easter island, guano and passenger pigeons.

All three of those things display varying degrees of 'diminished', 'depleted', and 'gone completely'. Which brings to the surface something important about renewable resources – when used, they are generally *used up*. They may be replaced by more of the same – if we show a little wisdom in their use – but they can certainly be depleted into extinction or non-existence.

In contrast, and with regard to the salient features of a finite resource - perhaps even the essence of the quality of *finiteness* - allow me to use the example of copper atoms. Both a single, humble instance of such, and the totality of copper atoms on earth. Though I have no advanced degrees (or, to be fair, any degrees at all) in chemistry, I'm led to believe that in normal circumstances in the terrestrial environment, a copper atom is indestructible, immutable and 'for ever'.

Another expression of this is that if there were a hundred units of copper atoms before there was any life on earth, there are the same hundred units now, and that number will persist



unchanged after all life on earth has ceased to be. The characteristic implied by calling a resource finite is that it is 'fixed' in quantity, and therefore un-depletable and in-exhaustible. Yet, even accepting this, there are intelligent people who argue the case that finite resources can 'run out' and I will come back to that in due course.

But for the moment I'm distracted by a large number of people shouting that "Coal is definitely being depleted!" and "How can you deny that oil is running out?!"

Well, yes it is, and no I don't – fossil fuels being things that perfectly exemplify how edge cases can mess up nice clean categories like *renewable* resources and *finite* ones. Take peat for example– a semi edge case that provides at least a theoretical escape from this little conundrum I've created for myself. Peat is a slowly renewing renewable resource that is laid down at the rate of about a millimetre a year. Easy to deplete, but renewable nonetheless. It is also halfway to being coal, at least in terms of its calorific values. Which makes it easier to see fossil fuels as having some of the characteristics of those depletable renewables, rather than things that are finite and immutable.

But whether or not you're happy to include fossil fuels in your category of the things that are at least theoretically renewable, Tim Worstall's book concerns itself solely with non-energy minerals and metals and the rest of this review will do the same.

## The Vastness Gambit

Throughout the *No Breakfast Fallacy* there are numerous references to how we're not going to run out of stuff for quite a while – really, for a long long while indeed. Some examples :-

Page 6 “*..we've been using the metals for decades and will be able to for millennia to come*”

Page 7 “*..there's nothing to worry about for tens of thousands of years at a minimum..*”

Page 16 “*We're not even going to have tight supply of anything for thousands of years..*”

Page 34 “*..there's not anything that we're going to run out of anytime soon.*”

Page 34 “*..there's not any likely shortage looming*”

Page 58 “*..there's no shortage of any metal or mineral that's likely to impact upon us or our descendants over any reasonable period of time*”

Go on Tim – tell us what you really think!

As it happens, I'm not in disagreement with him. But I don't think these assertions advance his cause as much as he'd like. If we, as a species, were as rational as some of us think we might be able to become, with a little bit of effort... this kind of approach to persuasion might work a treat. But we're not, and I think that for anyone whose motivating interest in resources is the worry that they're going to run out, it makes precisely no difference to argue that the point at which running out occurs is further away than they originally thought. The worried mind hears "Yes, we're going to run out.." and not the "But it won't be for ages and ages.."

Having said that, I was slightly misrepresenting Worstall by truncating one of the quotations.  
Back to Page 7 -

*"there's nothing to worry about for tens of thousands of years at a minimum, **more likely millions to billions**"*

[my bold]

This of course makes his vastness gambit somewhat stronger (if true..) but I remain unconvinced that many resource worriers will find that reassuring. The spectral catastrophe will still lurk somewhere in the dark, and every so often the sentiment will be expressed by the saying "We're still going to run out!"

## The Vastness Measured

One section of the book I really liked is the first appendix, called 'Mineral Estimates'. He discusses (not for the first time) appropriate ways to think about both mineral reserves and resources – what those terms *mean* to the people who use them professionally. Which I think are worth quoting at length :-

*"Mineral reserves are, as discussed earlier, the minerals that we know where they are, we have mapped and tested them, we know that we can extract them at current prices, with current technologies and make a profit by doing so. A rough and ready, if not quite exactly accurate, description is that these are the stock in trade, or deposits, at mines that are already being worked."*

Or, to use Worstall's own analogy – which he admits is rather tortured – minerals reserves are similar to the contents of the worlds fridges. Another comparison I've heard is that they're like the cars in a showroom, and mineral reserves no more represent 'all that there is left' than showroom cars represent the only cars that will ever be available.

He goes on to say that -

*“Mineral resources are the minerals that we know roughly where they are, have good reason to think we can mine at current prices, with current technology, and make a profit by doing so. But we’ve not gone through the expensive process of actually proving this.”*

It’s worth mentioning here (Worstell’s second appendix delves into the legal definitions surrounding this) that reserves are divided into *probable* and *proven*, and that resources are divided into *inferred*, *indicated*, and *measured*. And also that there is a fluidity between these divisions, strict as they are. Over time some part of a mountain can move from being an inferred resource to a measured one and then later to a probable or proven reserve.

However, in order to change, say, a measured resource into either a probable or a proven reserve you will need to spend heaps of money. You will need not just mining engineers and surveyors, but also lawyers and accountants by the truckload. The proving of a mine’s worth of copper, for instance, might cost you tens of millions of dollars. Which is the explanation for why there are only a few decades worth of each mineral in the ~~fridge~~ category of proven reserves.

The interesting part of the Mineral Estimates appendix is a table with a list of minerals and metals (from Aluminium to Zirconium) with quantities gleaned from the US Geological survey. It lists reserves, resources, annual production rates and how long those reserves and resources will last. More interestingly – because it really has sunk in that mineral reserves are, you know, just the working stock of already functioning mines? – are two other quantities for each of the 50 listed minerals: an estimate of the total amount in the earths crust and one thousandth of that.

This latter quantity is something Tim Worstell calls the ‘real resource’ and he admits to it being something quite fantastical. However, it’s for illustrative purposes only and I think in those terms it does a particularly splendid job. It’s saying “What if?” What if a mineral or a metal became so important to us that we mined one thousandth of the total amount in the earths crust? Just the richest deposits, the most concentrated ores, nearest the surface and furthest from environmentally sensitive areas? How long would that one thousandth part of the natural endowment of a resource last at current rates of production?

I think he hopes – as I would if I was on a mission to assuage people’s resource worries with the force of very big numbers – that the quantities and times will surprise his readers. I’ll mention three – aluminium, copper and lithium. I chose them not because the quantities and duration until exhaustion are exceptional compared with any of the other minerals, but because of their centrality to 21<sup>st</sup> century civilisation.

One thousandth of the total copper resources will keep us going for 600,000 years. The same proportion of lithium will be good for 70 million, and trusty aluminium will continue to be available for 270 million years.

Perhaps that’s one up to Worstell and the vastness gambit, but I’m not so sure. I will, I promise, come back to an alternative framing of the whole process, because he’s still seeing this as ‘running out’ .... just not for a very, very long time.

## A Triumvirate of Doubters

As the subtitle of *The No Breakfast Fallacy* implies, Worstall takes on the prognostications of the Club of Rome, specifically their 1972 book *The Limits to Growth*. He devotes a chapter to pointing out their misunderstanding of what people who know what they're talking about mean when they estimate mineral reserves. He has a chapter each for two others who also promulgate what might be called resource doomsterism. One is the financier Jeremy Grantham with his fears that industrial fertilizers will run out in a just a few centuries; the other is the contention in a 2007 issue of the *New Scientist* that humanity is about to encounter serious problems because -

*“...zinc could be used up by 2037, both indium and hafnium -which is increasingly important in computer chips - could be gone by 2017, and terbium – used to make the green phosphors in flourescent light bulbs – could run out before 2012”*

As you might imagine, Worstall thinks this is absurd, and I find it hard to disagree with him. Especially because when he was writing, Terbium was supposed to have run out three years ago. It's also an opportunity to give another flavour of his writing, which is direct, and uncompromising. A few examples :-

*“There is no excuse or explanation for the concatenation of piffle that the New Scientist served up..”*

*“Only someone possessed of the most absurd ignorance could possibly make such a howlingly incorrect prediction.”*

*“..the idea.. could only be advanced by people wallowing in their own purblind ignorance”*

*“..the claim..is the scientific equivalent of taking the short bus to school”*

Some of this I quite enjoy, probably because I share a few of his biases, and his sanguine view of resource availability [Yes, I might be repeating myself]. But there's something quite particular about the nature of these (and in general, all of his) criticisms of other people and their statements and beliefs. He exclusively criticises people for being wrong or incorrect on some issue – usually an economic one. He can be very extremely rude about people for their idiocy, misunderstandings and errors, and will freely call them cretins, morons and twats (at least on his blog). But he rarely, if ever, hints that he thinks people he disagrees with are evil or bad in any way – just possessed of too much stupidity and wrongness.

Maybe this is a distinction without a difference, but after I first noticed it, I realised it was significant for me. Perhaps it has something to do with a cultural and tribal contrast that was quite prominent in England a decade or two ago (and may not apply to current US red tribe/blue tribe conflicts). Those on the left would describe their opponents as the embodiment of evil, and those on the right would describe leftists as hopelessly wrong. My progressive friends still pepper their conversation with “I hate Tories”, whereas the conservative ones will say of Jeremy Corbyn that he's completely delusional.

It's probably not conducive to pleasant dialogue to call people twats and cretins, but somehow it rubs me up the wrong way a lot less than the 'repugnant lying scum' style of ad hominem so prevalent on the internet of today – amongst people of all political persuasions.

If there is some Englishness in Worstall's style of political criticism, there is plenty more in his sense of humour. And something unreconstructed that many people will find more problematic than I do -

*“Geologists are a hardy breed and the German sense of humour is notably robust. Getting a chuckle out of a group of German geologists usually requires a fart joke, or dropping your pants. And the men are worse of course.”*

I'm going to have to confess that when I first read this, I did in fact laugh a little. So perhaps I'm more of a bad person than I realised and that I should self-cancel or something?... I don't know, but it's fairly representative of Worstall's writing style and I'll leave any further judgement of it to you.

I wondered for a while why the Jeremy Grantham chapter (that there are only 3 centuries of phosphate and potassium reserves – we are therefore categorically doomed) was included in the book. Cynically, you might think it was there just to pad out something that was always going to end up as a pretty titchy little book. I mean, it's barely any longer than this review! But it also could have been because Worstall likes to point out the economic misunderstandings of the rich and famous. Indeed, it might be his overarching purpose in life.

More fairly, I think it's a useful platform from which Worstall can explain just why mineral reserves are the size they are – why they can't be much larger so everyone can relax a little more. Worstall claims (without providing much evidence) that Grantham fails to understand that money has a time-value; money invested in something that won't start making a profit for a hundred years will be better spent on something else which will provide a return much sooner. It strikes me as more likely that the billionaire financier understands the time-value of money well enough, but that he's like 'everyone else' and assumes that mineral reserves are something adjacent to 'all there is left'.

The Club of Rome's book *'The Limits to Growth'* is only two or three times the size of Worstall's pamphlet argument but has sold roughly a million times as many copies (and has therefore depleted the world's supply of ink by a similar ratio?) It has also been criticised in hundreds of ways by many hundreds of people, as well as having a large number of supporters.

Worstall, however, is quite admirable in his restraint, because I can imagine that every page of the environmentalists' tome has something to infuriate him, politically, philosophically or economically. He manages to keep his criticism focused very narrowly on the single misunderstanding of what constitutes a mineral reserve. And he assumes that this is sufficient to make their entire argument collapse into a heap of exhausted rhetoric.

Once he's pointed out this misunderstanding, he's – as usual – relatively direct about the consequences for their book.

*“The Club of Rome was wrong because they assumed at the start that there’s a relationship between mineral reserves and the mineral resources available to us. There isn’t, it’s a false assumption so they’re wrong.*

*So perhaps we should stop paying attention to them”*

## Surely You’re Joking Mr Worstall

If the thesis of *The No Breakfast Fallacy* is that humanity is most assuredly not going to run out of, or exhaust, any of the minerals we use for our modern society, what is going on with statements like these? :-

Page 6 *“The one thing that everyone does get right is that this generation is likely to run out of, consume all of, the minerals we currently have classified as mineral reserves”*

Page 43 *“..mineral reserves running out in a generation or two? So what?”*

I think there’s a strategic conceit going on here, a sort of pseudo-concession. He’s going to go on to say that reserves can be replenished – more times than you can shake a stick at. Therefore he’s happy to talk about our generation’s minerals being exhausted, depleted and consumed. But this gives an impression at odds with the reality of the published numbers and quantities of reserves. His characterisation would better fit the using up of the fuel in a vehicle’s fuel tank – it starts off full, over a period of time becomes depleted, and then at a particular juncture gets filled up again.

Perhaps I’m being overly critical here – he’s merely making a rhetorical choice, after all. But a look at the US Geological Survey’s published estimates of mineral reserves do not show the quantities going down. Not over a decade, a generation or a century. There is no period of emptying followed by an instant of replenishment – the process is more like a steady topping up. To add yet one more analogy to the mix, it is as if you’re in a supermarket buying baked beans (Heinz please – why compromise?) and moments after you’ve added a tin to your basket, a friendly shop assistant adds another one to the shelf, keeping the size of the reserves as they were.

## Abundance and Scarcity

I agree with Tim Worstall that resource worries are pretty much always based on a misunderstanding of what the word *reserves* means (to those who makes the calculations about the quantities involved). And he makes a convincing case that supplies of all the finite minerals and metals will be adequate for our needs for a longer time than it makes sense to even think about. But – and I did hint that there would be a ‘but’ - as radically different as his and the

resource-worried positions appear to be, I think they both misconceive the nature of the process that is occurring as we mine those metals and minerals.

I'll start with the most difficult to argue against position, because after that everything will be easy peasy and a walk in the park. The contention I want to dispute – and offered by some of the wisest people I know – is that “There are theoretical reasons why resources have to run out eventually.” I respectfully beg to differ, but to begin with I'll do that Worstellian thing and appear to concede the whole argument. I'll admit that it's undeniably true that there's a point in time where resources – any of them/all of them – are not available. There's none to be had at any price. There are no reserves at all and nothing is being produced whatsoever.

My difference of opinion is about the *when* of this point in time – resource depletionists think it is at some point in the future. I'm convinced it is in the past.

Let's take aluminium as an example and use some measures of resource abundance to see whether it makes sense to say that it is heading towards scarcity.... or away from it. Are we running out, or running in?

We could use four measures of the abundance or scarcity of a resource like aluminium. Firstly the familiar estimates of proven reserves – the greater the reserves, the greater the abundance. Secondly the amount being produced (per capita to control for population growth). Thirdly, and only a slight variation on the second – the amount in use or circulation. And lastly the price of the resource – the less you have to pay for something the more abundant it is.

By these measures the abundance of aluminium in 1824 was zero – it effectively did not exist. No reserves, no production, none in circulation and no amount of money would have bought you even an ounce of it. The metal as we know it was innovated into existence in that year. And by all four measures it has been getting ever more abundant since then. More reserves, more production, more in use and cheaper over time. All the while, the number of atoms of aluminium on earth has remained unaltered (“No aluminium has been harmed in the making of this ~~film~~ tin can”) The only change has been the amount being put to use by human beings.

Coming from a position of absolute scarcity it is inevitable that the changes in abundance will have been dramatic. Even since the mid 19<sup>th</sup> century (when aluminium was more expensive than gold) it has become four orders of magnitude cheaper and reserves have increased by similar numbers. By 1893 it was sufficiently abundant that it could be used to make the staggeringly beautiful statue that graces the corner of Piccadilly Circus in London, facing Shaftesbury Avenue, and often mistaken to be of the god Eros.

If aluminium is perhaps a recent and freakish exception to the rule, what about the other end of the spectrum – the metal that we've used for the longest period of time – copper? We've been utilizing it for the best part of 10,000 years, and as it happens, the same trends have occurred as have with aluminium. We have progressed from ultimate scarcity to increasing availability, without any change of the amount of copper on earth at all. Ten millennia of increasing abundance.

If you're familiar with the story of 'Otzi the iceman', the mummified remains of whom were uncovered in 1991, you'll know that this character is assumed by archaeologists to be wealthy and powerful primarily on account of his possession of a copper axe. So, 5,000 years ago when Otzi was shuffling around in his deerskin snowshoes, copper was still extremely scarce.

For comparison, in my rural woodland world, axes for chopping firewood come in two sizes – 88 ounces and 130 ounces (though thanks to Bonaparte they've all been kilo-ed and gramme-ed). Otzi's axe, which marked him out as such a wealthy individual, weighed precisely 6 ounces. It was less of an axe and more of a small chisel strapped to a stick, but scarcity keeps prices high! The same amount of copper today would signal neither wealth nor power, and would cost you less than the price of a Big Mac (no soda, no fries). By all four measures of abundance, and just like aluminium, copper appears to be not running out, but running in.

## More Abundance and Less Scarcity

If I were Tim Worstall and wanted to pad out my pamphlet sized argument about the non-running out of mineral resources, I would use an example from Wikipedia's page on peak copper. Environmental analyst (and resources expert..) Lester Brown is quoted as suggesting in 2008 that copper..

*“might run out within 25 years based on what he considered a reasonable extrapolation of 2% growth per year”*

I've been tempted to edit that entry in Wikipedia a number of times but a) it would get reverted immediately on the grounds that Lester Brown is very famous and he really did say that, and b) it's cringeworthy, especially from someone who spent their whole working life as a resource analyst, but it is also a truly wonderful testament to the ubiquity of resource misunderstanding.

About ten seconds and some fairly basic Google-fu will find you on a US geological survey page detailing that the production of Copper increased by a factor of 40 over the 20<sup>th</sup> century (with the cost falling) But it is equally easy to find the data showing that copper reserves increased by a factor of (wait for it..) 40 as well! Which is no surprise to anyone who is familiar with some of the analogies about what sort of thing is a mineral reserve. Another analogy is that they are the portal through which we bring the resources into use, and the size of the portal (the reserves) grows as we change the rate at which we bring the resources into circulation.

But what about the amount that's left in the ground? Doesn't that become diminished? Well, one response is the vastness argument, which says that the percentage of any mineral resource that remains in the ground will always be 100 (to the nearest whole number). And that argument may well be sufficient for many (including Tim Worstall). But the reason I don't think the reserves of anything can even theoretically run out is because we can never get to one percent of resources in circulation. And this has nothing to do with supplies or availability. Imagine what would have to be the case if one percent of the mineral resources of the earth were in use.



Just to be clear, we firstly need to know what proportion of the earth's crust is resources. Take a cubic hectare from under a random piece of the earth's surface. Let's choose... for the sake of an example... Lords cricket ground in London (why not – if it's good enough for Arthur Dent and Ford Prefect?). We'll have a couple of million tons of rock to play with. If we processed it and sold all the resources on the global market, what proportion would be left?

Zero percent, of course (to a whole number again, for the nitpickers among us, myself included). And that's true for every cubic hectare of the earth. It's resources all the way down! And the reason we don't immediately rush to Lords and start digging, is because there's more money to be made digging in other places where the concentrations of particular resources are higher. As well as where there are fewer cricket matches.

So, if we did in fact bring one percent of the earth's resources into use, how would the surface of the planet look? Well, the earth's crust is about ten miles thick, so one percent of that will be....about..... 500 feet. That's 500 feet of solid resources (I guess the gases will be in canisters) in amongst us and above our heads. Right round the earth, no gaps. How could this be? How would sunlight get in? How would we move about?

This isn't exactly a vastness argument, it's more of an inconceivability one. To me that is an inconceivable vision. It is also something which I'll take as a rebuttal to the theory – that we must run out of stuff *eventually* - most days of the week.

Perhaps another way of sifting through this confusion is to notice that the language appropriate to one kind of thing can create widespread confusion if misapplied to another. It is fair to say that the bacon in your fridge can be used up. It can be *consumed* (and not just economically so). With a sufficient lack of care we can run out of it in our fridges, on earth and in the universe at large.

But to say that a copper atom can be used up is to misunderstand its finiteness. All we can do is move it from one place to another, and make more or less use of it over time. This is an open-ended process; there is no fundamental limit to the increasing efficiency with which we use copper atoms or any of the other materials available to us. Tim Worstall is right that we're not going to be running out of minerals and metals anytime soon. Or even anytime not very soon. But the situation is better than even he imagines - the running *in* of physical resources is unbounded, and in human terms, forever.

## Abundant Addenda

1. For technical reasons there are no links in this review. Typing the words 'No Breakfast Fallacy' into Google will take you almost directly to a pdf copy of Tim Worstall's book if that interests you. It really is very short.
2. 'The Limits to Growth' can be read for free through the same process.

3. If I were to recommend one article on the subject of resource (specifically copper) depletion, it would be the blog post 'A Rosy View from the Patisserie' by Geoff Batt on his old blog Rocky Subjects. It is informed, intelligent and beautifully written. It is also a lot less shouty and much more nuanced than the writing of either Mr Worstall or myself.

# The Open Society and its Enemies by Karl Popper

In October 1958, the journalist and author Bryan Magee attended the presidential address at the Aristotelian Society in London, a talk “attended almost entirely by professional philosophers, many of them well known”. The speaker was the philosopher Karl Popper, and his topic was the idea that knowledge grows from the feedback process of criticism, something which originated with the pre-Socratic Greeks from Thales onwards. Magee watched with incredulous disbelief as the subsequent discussion missed Popper’s central point, focussing on the minutiae of his historical references, stating “instead of presenting his revolutionary idea head-on, he had presented it indirectly, in the form of a historical claim about the pre-Socratics”, and that, “he was presenting his ideas in a way that almost ensured they would be misunderstood”.

This is certainly the case with his seminal work ‘The Open Society and its Enemies’, written in New Zealand during World War II, as Popper (who was of Jewish descent) fled the Nazis from his home of Vienna. A polemic against totalitarianism, Popper attacks the ideas of Plato, Hegel and Marx, attributing to them the intellectual roots of Stalinism and Nazism. Much of the subsequent criticism of The Open Society is a wounded defence of those great thinkers, rather than any true assessment of Popper’s thesis.

Primitive societies according to Popper, would not distinguish between natural laws and customs, since the two appeared entwined. Reliance on unexplainable natural phenomena such as weather gave importance to custom and passed-down, often mystical wisdom. Under such conditions, individual accountability for moral decisions and individual autonomy was low, and generally there was little questioning for the way in which things were done. This is what Popper calls a Closed Society. As populations grew, a Closed Society might establish ‘daughter settlements’ to handle increased populations, but geographical distance, and (especially) a tendency to trade would lead to customs being challenged. This describes Athens around 400BC, as a limited form of democracy emerged, and the pre-Socratic philosophers began to question all superstitions - the first moves toward an Open Society. Contrast this with Sparta at the time, who’s leaders sought to arrest change by shutting down all outsider or democratic influence, to avoid all trade, and to limit the size of the state. It is important to note that the fear of change (and its possible personal consequences) on behalf of the tribal leaders motivated such a regime.

It is critical to point out that Popper's objection to Closed Societies is epistemological rather than strictly political, and can be traced to his Philosophy of Science, his true passion. Arguably Popper's greatest achievement was his proposed solution to the problem of induction, with his idea of falsification. Briefly, sceptical philosophers like Hume had shown that from a logical standpoint, no quantity of confirmatory evidence can actually prove the truth of a statement. Seeing the sun rise for 10,000 consecutive days, in no way proves that it will rise on day 10,001. This is the riddle of induction. As Bertrand Russell stated, a chicken may believe for 363 days that his master is a generous soul who feeds him every day, and on day 364 it is Christmas, and the generous master chops off his head. Popper's solution is that we can't ever prove anything to be true, but we can falsify a statement. The statement 'all swans are white' can never be proven by seeing additional white swans, but a single black swan will falsify the theory. If we create falsifiable statements and theories that tell us more about the world than their predecessors, our knowledge will progress.

By this logic, knowledge doesn't grow from passing down wisdom so much as the creative pursuit of falsifiable theories that better explain the world. Einstein's relativity didn't build on Newton's laws so much as explode them. Applying this 'meta-rule' to the political sphere, Popper attacks Plato, Hegel and Marx's 'Historicism' - their propensity to look for patterns in social science and use them to ascribe predictive 'laws'. The method Popper calls Historicism gives its creator the superficial likeness to science - but since these theories can never be refuted (or falsified), there can be no actual progress in using them. Instead he prescribes what is a simple definition of democracy. What states ought to do argues Popper, is to create a situation in which a bad leader can be removed without violence. The question should never be 'who should rule', but instead 'how can we organise political institutions so that bad or incompetent rulers cannot do too much damage?'.

First in Popper's crosshairs is Plato, student of Socrates and teacher to Aristotle, together the three most important ancient philosophers. Popper describes Plato as the greatest ever philosopher, but spares him no scorn. At face value, it is a reasonably simple matter to attribute some blame to Plato for the rise of totalitarianism. In *The Republic*, (the most widely read book of the most influential philosopher of all time), Plato advocates a society in which the Philosopher-King lies to the people, creating a fake ballot to favour the breeding of the 'warrior class', and creating a myth to justify segregation on racial lines. He states that all children over the age of 10 are to leave the city to create a 'clean canvas', and babies within the ruling class are to be separated

from their parents at birth (with all adults living communally). Popper compares this 'blank canvas' radicalism with both Lenin and Hitler's various acts to 'purify, purge, expel, banish and kill.' This he attributes to the idea of Utopian Engineering. It may seem natural to choose an ideal end goal for which a nation should aspire, and make all effort to reach that goal. The problem argues Popper, is that we are not aware in advance of how to reach that goal, and that radical reordering of society without such knowledge brings little benefit at enormous cost. Radical social engineering requires a powerful dictator, who must be deaf to the criticism of the disruption caused en route to the final goal; but since no complex society can know a priori how best to reach that goal, in practice such a leader will be deaf also to genuine, valid critical feedback.

In short, the problem with a Utopian state is that one needs both a rational way to determine the best end goal, and a rational way to determine how to get there. Since neither exists, the direction will be chosen by some form of intuition or opinion - and so the chosen path is decided by power, and ultimately violence. Instead of Utopian Engineering, society should, argues Popper, practice Piecemeal Engineering. The trial of setting up an insurance company, or building a school or new roundabout will give us information as to its usefulness at little cost. We will explore this idea further when discussing Marx.

Can we judge Plato by the standards of today, when he grew up in a society in which slavery was commonplace? Would it be reasonable to condemn Martin Luther King as an evil tyrant in fifty years time for eating red meat if, by then it was considered a despicable practice? The context in which Popper places Plato's later writings indicate that we can. Socrates never wrote anything - we only know him from the writings of Plato (and some other of his contemporaries). Popper contrasts Socrates' intellectual humility with Plato's insistence on a ruling elite of philosopher kings, and calls it a betrayal of his teacher. In fact, he accuses Plato of outright deception in his writings, deceptions which have had lasting impact on the history of thought.

In his treatment of the concept of 'justice', Plato performs, according to Popper, a sleight of hand. By conflating individualism with selfishness, and collectivism with altruism, he has created an intellectual tool to be used by future despots. Next, by asserting the importance of sovereignty - that 'the best' or 'wisest' should rule, Plato is openly espousing anti-democratic sentiments.

Plato pointed out the 'paradox of freedom and democracy'. If it is the will of the people to appoint a strongman dictator, democracy will end itself. This was apt in the 1940s, as Hitler's ascent was partially democratic. Popper counters that all forms of sovereignty

are paradoxical - if we appoint 'the wisest' in their wisdom they could wish to appoint 'the strongest' or vice versa.

Further, we can avoid the paradox of democracy by identifying 'government that can be gotten rid of without bloodshed' as our main ideal. This distinction is Popper's definition of democracy versus tyranny. Simply, a government that can be peacefully gotten rid of. Therefore democracy should not be seen as 'rule of the people' so much as the choice of removable government over unremovable tyranny (however wise or benevolent it may be).

It is the distaste for change, according to Popper, that drives Plato's advocacy of an unchanging, 'perfect' state. Socrates was a friend to democracy by his constant challenging and questioning of it - a service that would be outlawed in Plato's Republic. Because he is both a masterful essayist and unparalleled thinker, his historicism and utopianism have influenced countless others. Our inability to know the best future path can tempt us to long for a 'philosopher king' to absolve us of the responsibility of difficult decisions. "If our civilisation is to survive", says Popper, "we must break with the habit of deference to great men."

The second major target of Popper's scorn is the German philosopher, George Wilhelm Friedrich Hegel. Writing from exile in New Zealand, it is claimed that Popper's access to Hegel's material was limited; indeed many Hegel supporters state that Popper's attacks are wide of the mark - both unfounded and unfair. From reading the book the link from Hegel's philosophy to totalitarianism isn't as clear cut as that of Marx and Plato. Much of Popper's critique of Hegel is a personal attack. Briefly, Popper charges that Hegel has taken Kant's extremely dense (yet coherent) writing style, and taken it yet further, deliberately using unintelligibility as a mask for charlatanism. Further, he contends that his pedestal is used to prop up Friedrich William III, his employer. Hegel's assertion of the absolute moral authority of the state is, contends Popper, a sinister and intellectually corrupt act intended to prop up an undemocratic state. This is not unfounded.

The great Philosopher Schopenhauer, who knew Hegel personally, stated "Hegel, installed from above, by the powers that be, as the certified Great Philosopher, was a flat-headed, insipid, nauseating, illiterate charlatan, who reached the pinnacle of audacity in scribbling together and dishing up the craziest mystifying nonsense." Indeed, it is worth quoting Popper's quote from Hegel in full, to illustrate quite how absurd his writings could be. Says Hegel, "Sound is the change in the specific condition of segregation of the material parts, and in the negation of this position;-merely an abstract or an ideal ideality, as it were, of that specification. But this change, accordingly, is itself

immediately the negation of the material specific subsistence; which is, therefore, real identity of specific gravity and cohesion, i.e.-heat. The heating up of sounding bodies, just as of beaten or rubbed ones, is the appearance of heat, originating conceptually together with sound.” Popper calls this gibberish, and, the language barrier notwithstanding, it is difficult to disagree.

The real question is whether or not this is important? If Hegel is a liar and a clown, whose success is based upon corruption, how has this contributed to the desperate state in which the world finds itself by the early 1940s? Again, we must return to epistemology. It is Hegel’s method of historicism to claim the central importance of the state that matters. If those sitting on an intellectual pedestal can claim to see historical trends, trends that predict future events, these can become a powerful force in and of themselves. An ‘intellectual fifth column of the state’, as Popper says. For if such predictions are unfalsifiable, then they can predict one direction or another, and can be corrupted. His idea of ‘progress’ of society, his dialectic argument that we are advancing, has an implicit moral implication - might is right. If a historicist like Plato believes in a starting state of perfection (his ‘Forms’), then what was past is what’s good, if Hegel’s historicist idea of ongoing progress is believed, then what is now is what is good, and if Marx’s historicist prediction of the future state of society is believed, then what is believed to be next is what is good. Popper’s point, beyond his firm belief in the inability of anyone to predict future trends (or even in their existence), is that our ethics cannot stem from the state of play - for why would what is right and wrong necessarily coincide with what is?

Finally, Popper turns to Marx. At a time when the Soviet Union was on the side of the allies, Popper dedicates at least a third of the book to comprehensively dismantling the ideas both of Karl Marx, and those he calls ‘Vulgar Marxists’. Popper’s method is particularly effective - he first lays out the opponent’s arguments in their strongest possible form, often improving them along the way, before refuting those arguments absolutely. There is no personal criticism of Marx in the way there was for Hegel - in fact Popper praises Marx for both his intentions and his remarkable insight.

Significant space in the book is spent addressing and critiquing Marx’s specific labour theory of value, a derivation of 19th century Ricardian economics. To go into detail here would be to double the length of the review, and Popper himself states that if the theory were wrong (as he believes it to be), it still doesn’t invalidate Marx’s theory. Therefore we will omit that discussion.

In *Capital*, Marx isn’t blaming psychology for the world’s problems; his is an institutional

sociology - including his discussion of class war. This is in opposition to 'vulgar marxists' who assign sinister motives to groups. Marx's argument is that the economic system itself traps both classes - those who live in forced servitude, and capitalists, who, trapped by competition, must always exploit the worker or be driven out of business. Marx says we are not free unless we escape the productive process- that our metabolism comes before our free mind (an inversion of Hegel - 'the kingdom of freedom versus the kingdom of necessity'). However, the way to escape is to use other people's productivity for our own ends. And in this way there is no escape- since both oppressor and oppressed are trapped in class struggle. Therefore any period in history is defined by the economic system - 'feudalism', 'capitalism' etc. Individuals - capitalists and workers, cannot change the system consciously. The legal system according to Marx, is the apparatus by which one class uses force to oppress another, as part of the superstructure of economic society. The upshot is that since the means of production determine the state of play, political systems are impotent, or superficial. When they change they are a consequence not a cause of real change. Popper contends that whilst studying institutional class struggle is useful and worthwhile, it cannot be the sole explanation for the state of the world, for example, when emperors fought Popes in medieval history, this cannot be interpreted as oppressors fighting the oppressed.

There are two key mistakes worth discussing here. Firstly, in Marx's construction, political power is the third consideration, ranking after the evolution of machinery (or today, technology), and economic power. From the vantage point of the industrial revolution, this may have seemed natural to Marx. Even today, one can convincingly argue that money allows one to buy power. But in our dissatisfaction with the inequalities and corruption of the modern world, we cannot lose sight of the essential truth that no wealth can protect against political power in the final reckoning. The supposed power of the jews stemming from economic power was of no use in Germany under Nazi rule, nor the middle classes in general under Mao or Stalin. Even today the billionaire oligarchs of modern Russia are subject to Putin's demands at pain of imprisonment. Centrally, Marxists have failed to sufficiently appreciate that power must be constrained in all its forms in order to avoid exploitation.

Leading on from this is the second critique, where Popper is most convincing. He has much sympathy, he says, with Marx's strong opposition to "the most shameless and cruel exploitation", which was "cynically defended by hypocritical apologists who appealed to the principle of human freedom". In 1863 when Marx was writing Capital, he saw in England 7-year-olds performing 15 hours of labour a day, women dying after working 26 hours of straight work, and countless other horrors of exploitation, "defended not only by professional economists, but by churchmen". It is no wonder says Popper,



that Marx had no faith in either liberalism or parliamentary democracy. Popper's genius is in invoking the Paradox Of Freedom. Freedom he says, defeats itself, if it gives the freedom of a strong man to bully a weak one. This must be extended to economic matters. I cannot use my economic power to put you into servitude. Your starvation does not give me licence to unrestrictedly use your labour. This cannot be remedied by the state 'withering away' as Marx put it, but by legal means, as with protection from rape or murder.

This is critical, because it puts front and centre the opposing doctrines of Popper's 'Piecemeal Engineering' with Marx's 'Utopian Engineering'. Whilst Marx predicted an increase in exploitation and misery as capitalism developed, the opposite, in fact occurred. The age at which one can work has been progressively raised, whilst the hours one can legally work progressively reduced. We have gone further still. All workers in England now have access to universal healthcare, education, disability and unemployment benefits and emergency housing. None of this required the overthrow of capitalism. The progress made argues Popper, is a result of democracies asking not 'Who Should Rule?' (remember this was Plato's question, and in another form Marx's, the answer being 'The Proletariat'), but 'how can we remove without violence bad rulers?'. By adhering to Popper's question rather than Marx's, progress has been slow but real. In fact, of the ten demands made in Marx & Engels' Communist Manifesto, many don't seem so revolutionary today; a heavy progressive or graduated income tax, central control of the state of the means of communication & transport, free education for all children, abolition of child's factory labour. Looked at from the opposite point of view, Popper contends today that Laissez Faire has vanished from the face of the earth. Whilst adherents of Reaganism or Thatcherism may have called their brand of neo-liberal economics 'laissez faire', in truth those programmes bore no resemblance to the horrors of Marx's time.

One may conclude from subsequent events then, that the horrific aspects of capitalism witnessed by Marx were in fact a bug, not a feature. This is true even of the less controversial observations, such as the tendency to increased concentration in firm size (solved through legislation), and the worsening of trade cycles (solved by Keynesian and monetarist interventionism).

A perfunctory search for reviews of this book turns up bile from libertarians, disappointed that Popper was not an outright anti-government dogmatist in the manner of his friend Hayek, from philosophers for his attack not just on Plato and Hegel but on the profession's obsession with language when there are real problems to solve, and

from Marxists and communists of all kind, for the destruction of their entire thought system. Despite his reputation as a fearsome and often dogmatic intellectual, and especially despite the ferocity with which Popper attacks deceased and revered intellectuals, the most striking aspect of *The Open Society and its Enemies* is it's...reasonableness. Popper was undoubtedly a genius, a genius who found himself an undeserving fugitive from the worst evil humankind had to offer. He took all his intellectual gifts, and argued with almighty power in favour of a calm, gradual, considered, reasonableness. He recognised the gains to be had in intellectual life from grand system-building prophets, and the damage such men could do. And he fought against them with logic and reason and truth. A book as prescient and important today as it was in 1944, for anyone with an interest in the wellbeing of society, this is the top of the list of must-reads.

# The Oxford Handbook of 4E Cognition by Albert Newen, Leon de Bruin, and Shaun Gallagher

**Abstract:** The “4E” approach to cognition argues that cognition does not occur solely in the head, but is also embodied, embedded, enacted, or extended by way of extra-cranial processes and structures. Though very much in vogue, 4E cognition has received relatively few critical evaluations. This review performs such an evaluation with reference to The Oxford Handbook of 4E Cognition, which represents the most recent research in this area

## Book under review

Albert Newen, Leon de Bruin, and Shaun Gallagher. 2018. The Oxford Handbook of 4E Cognition. Oxford: Oxford University Press. 960 pages. Hardcover \$165.00.

## Review

We have categories for three types of cognition: organic, artificial, and supernatural. Most organic cognition that we know of is carried out by fish and arthropods, which together make up eighty-five percent of Earth’s non-vegetable biomass (Bar-On, Phillips, and Milo 2018). Artificial cognition is relatively new, but already able to outperform human judgment in a range of specialized tasks (Goodfellow, Bengio, and Courville 2016). We have no evidence whatsoever that supernatural cognition exists, and it would probably do us a great deal of good to stop looking for it.

If I make these points, it is because any appreciation of how cognition functions must begin with an appreciation of what counts as cognition in the first place. One recent volume purports to do this, giving a wide selection of essays by well-known and newer voices in the 4E cognition paradigm. In different ways, these essays respond (or fail to respond) to the issues that attend thinking about 4E cognition, and thus offer a useful point of entry into the field. But before engaging with the essays themselves, I offer some thoughts on the wider intellectual project that they contribute to.

The four E’s of 4E cognition initialize its central claim: cognition does not occur exclusively inside the head, but is variously embodied, embedded, enacted, or extended by way of extra-cranial processes and structures (Rowlands 2010a). This is intended to contrast with traditional views, where internal representations of states of affairs are symbolically manipulated so as to output knowledge or action. Consider my planning of my monthly finances using an electronic calculator. Depending on the 4E theorist, this process can be described as embedded cognition (it facilitates thinking by causally exploiting an object in the environment) or extended cognition (the calculator constitutes part of my cognitive apparatus). However, the action also enacts a world: by budgeting

for the future, certain items in the environment are disclosed to me as affordable or not affordable. Finally, my physical and cultural embodiment as a human being shapes my cognitive processes, such that I am most comfortable using decimal arithmetic (I have ten fingers) and will avoid thinking with degrees of precision greater than is practically useful (enculturated knowledge). Inevitably, the 4 E's shade into one another, and authors differ in how they define them. Nevertheless, most agree that they constitute a form of dynamic coupling, where the brain-body-world interaction links the three parts into an autonomous, self-regulating system.

As a historical phenomenon, the 4E paradigm draws on several precursors. Early work on externalist accounts of truth showed that it is possible to construct a semantics in which the meaning of a proposition is not solely determined by the intentions and beliefs of an agent, but also by external (and possibly unknown) facts in the environment—the famous “twin Earth” thought experiments of Hilary Putnam (Putnam 1973). In psychology, work by James J. Gibson on perception and affordances (Gibson 1979), Lev Vygotsky on socialization, thought, and language (Vygotsky 1962), and George Lakoff and Mark Johnson on conceptual metaphors (Lakoff and Johnson 1970), demonstrated how cognition is unavoidably shaped by the somatic and cultural environment it occurs in. But perhaps the most fateful encounter in the development of the 4E paradigm came with the blending of the continental phenomenological tradition (most importantly the work of Maurice Merleau-Ponty [Merleau-Ponty 2002; 1996]) with the methods of contemporary neuroscience. The ambition here was to reconcile “the body as a lived, experiential structure and the body as the context or milieu of cognitive mechanisms” (Varela, Thomson, and Rosch 1991, xvi). In a program that was to profoundly shape subsequent research in 4E cognition, this involved complementing third-person knowledge of specific brain states with a phenomenologically informed, first-person perspective on what it's like to experience these states. Most recently, this has been taken up in relation to social cognition, where 4E theorists argue that internal representations of social scenarios are not necessary for us to successfully engage with these scenarios—instead, they are processed directly by neurological mechanisms or delivered by cultural prostheses such as narrative (Gallagher 2012; Popova 2014; Hutto 2009; Carney, Włodarski, and Dunbar 2014).

Relative to rival models of cognition, there can be no doubt that the 4E paradigm represents a welcome turn to how much of cognition actually occurs. Cognition is always directed towards some end, and as these ends are usually practical, it is reasonable to suppose that the means by which we most often pursue them—our bodies—are integral to the understanding of cognition. There are also good a priori grounds for thinking that something like 4E cognition makes evolutionary sense. Cognition is physiologically expensive, so if a regularity in the environment can be depended on to perform computational or inferential work, any organism that exploits this regularity can divert more resources into successful reproduction. It is therefore no surprise to find evidence in support of the 4E account. One area of particular interest is in the psychology of perception, where action-based theories suggest that the environment is disclosed through active probing—as opposed to the passive registration and subsequent synthesis of sense data (Noë 2004; O'Regan and Noë 2001; O'Regan,

Myin, and Noë 2006; Bogacz 2017; Adams, Perrinet, and Friston 2012). More speculatively, it is also true that the entire apparatus of culture can, in many ways, be thought of as the displacement of challenging cognitive operations into material and symbolic artefacts in the environment (Malafouris and Renfrew 2010; Carney, Robertson, and Dávid-Barrett 2019; Hutto 2009; Gallagher 2011). Indeed, if “environment” is also understood to include the body, good empirical evidence has recently emerged that even abstract, culturally mediated concepts like “justice” and “essence” are encoded using perceptual and motor schemes (Connell, Lynott, and Banks 2018; Lynott et al. 2019). And we need no evidence at all to know that the world matters to us without us needing to reflect on the fact—a facial expression is immediately visible as sad or happy; we do not need to deduce this from evidence (Gallagher 2008). In this sense, a key value of 4E approaches is to make manifest those preconditions of thought that are invisible precisely because they are so fundamental.

And yet, one can acknowledge the value of the 4E paradigm, and even contribute to the literature, without subscribing to the revolutionary zeal that too often attends it. Cognitive science is not, on the whole, a discipline known for its febrility; writing on 4E cognition—especially when delivered by authors not involved in experimental research—is the exception to this. If there is a trademark register in the field, it is the angsty ponderousness that often attends politics passing itself off as science. What these politics might be is signaled by the rhetorical overemphasis on the material and the somatic that attends so many expositions. “The zone of the carcass and the knacker” (Adorno 2002, 117) has always been the favorite starting point of progressive thinkers, and in thematizing it, the 4E paradigm signals a curious affinity with the various iterations of Marxism, feminism, phenomenology, and queer theory that make the same move. Indeed, this alignment points to one reason why the 4E paradigm is making inroads into humanities scholarship when fields like evolutionary psychology are daubed with the mark of Cain. Evolutionary psychology foregrounds reproduction, and Mom and Dad are the least sexy people around; 4E is about bodies, and bodies are very sexy indeed. Combine a body with a historical wrong—the repression of its role in cognition—and we’re already halfway to the Finland Station.

If I exaggerate here, I do so to make visible the extent to which the arguments presented by 4E theorists are sometimes not arguments at all, but rhetorical appeals to the reader. We are told, with a straight face, that “when the phone rings, I hear John’s voice; I do not hear a set of electronic sounds and infer that John is the cause of these noises” (Moran 2017, 32)—as if the italics do anything other than change the typography. Inferring John’s voice from the electronic sounds is precisely what comprises hearing; inference is built into the process. I quote this example not because it is egregious, but because it is typical. Throughout 4E writing, the aversion to anything that smacks of inference introduces a dualism that would be untenable were it applied to physical processes. When I hear the phone ringing, do I detect from the vibration of air molecules on my eardrum that a socially agreed signal has supervened? The answer must be yes, unless I wish to deny materialism (which is to say, affirm magic). That I do not consciously experience the calculation doesn’t mean it isn’t being performed. This is

a common move in the field, where the forceful assertion of phenomenal experience is presented as an argument in its own right. We are already familiar with this mode of argumentation—it's called literature—but where the errors of literature are paid for in pleasure, we are rarely so lucky with 4E cognition.

A more serious problem lies in how the undeclared politics of the 4E paradigm subtly fore-stalls criticism. To query 4E cognition is to be somehow gauche: bodies are where it's at, and you must surely wear socks with your sandals if you think otherwise. But the fact remains that there are foundational issues with the 4E paradigm that are too easily ignored or dismissed by its protagonists. Here are some of the more pertinent ones:

*The hidden processing objection:* 4E theorists correctly claim that our bodies deliver us a world already freighted with significance; we do not usually consciously establish our interests with calculation and inference (at least with respect to perception and low-level action). But this is only true because the calculation and inference have already been carried out by evolutionary selection and hardcoded into our genome. To the extent that 4E theorists respond to this, they do so by offloading the relevant inferential activities onto some occult neurological process that implements them without the need to invoke symbolic intermediation. Gallagher (2008), for example, argues that social cognition is “directly” implemented by mirror neurons. In computing terms, this is analogous to saying that no computation occurs when graphics rendering is performed by a GPU (a built-for-purpose graphics processing unit) instead of a CPU (the general-purpose processing core), because the GPU just “delivers” the result to the CPU. Surely, such a view impoverishes our knowledge of cognition instead of enriching it. It does not matter how, when, or where the inferential activity occurs; if it is a necessary concomitant of a cognitive process, then any worthwhile approach to cognition is obliged to account for it.

*The intentionality objection:* No one demurs when we say that both human beings and giant squid practice cognition—despite *Architeuthis* and *Homo sapiens* inhabiting vastly different ecological niches. This is because we see different forms of cognition as having common features, like intentional directedness, that distinguish them from other organic processes. Thus, to say that cognition is shaped by the body that sustains it can be true without at the same time being useful. Human thoughts and squid thoughts are cognate in ways that human thoughts and squid heartbeats are not, because thoughts, unlike heartbeats, are “about” something. Moreover, as soon as we inquire into what this aboutness might mean as a phenomenon distinct from its token implementations, we find that bodies stop mattering very much at all. If thoughts are expressible and communicable, it can only be because they have some feature by virtue of which they pick out stable features of the world that can be accessed in an intersubjective way. We do not know for sure how this occurs—intentionality remains a riddle—but by far the best worked out candidate is the computational approach. Formulations like the Turing-Church thesis give clear definitions of what it means for a function to be computable, and thereby provide a formal model for cognition. In this view, the intentionality of thought is explained as a functional mapping from inputs to outputs, and the communicability of thoughts is volunteered as an isomorphism between compositions of mappings. Is this view correct? Perhaps it is and perhaps it isn't, but there is no rival

view that even remotely approaches the computational perspective with respect to theoretical sophistication and consilience with the rest of knowledge. Thus, to the extent we are interested in cognition in general, it may be that the 4E paradigm misleads us into emphasizing its incidental features over its essential ones.

*The duplicity objection:* Human beings habitually practice deception. Some authors even claim that this is responsible for the evolution of the cortex in the first place (Humphrey 1976; Hill and Dunbar 2003). If so, there is risk in naively accepting the interpretations of our social environment that our bodies deliver to us. Instead, we would be better advised to evaluate, in a Bayesian manner, the plausibility of these interpretations against contextual evidence, prior knowledge, and counterfactual models. With friends and family, such processes may be damped or suspended; with strangers they will be amplified. In between the two are those mechanisms of social association that, to greater or lesser degrees, allow us to identify strangers as allies or proxy kin (Dávid-Barrett and Dunbar 2017; Whitehouse 2000; Cohen 2012; Launay and Dunbar 2015)—even if these must then be coupled with mechanisms that prevent free-riding. To this, the 4E theorist will make the reasonable objection that we have pre-reflective intuitions for lie detection that help us negotiate tricky social environments—and indeed, it is true that there is a class of individuals in whom this skill is particularly well developed (Ekman, O’Sullivan, and Frank 1999; Ekman and O’Sullivan 1991). However, social life is only ever partially conducted through face-to-face interactions; even in the Palaeolithic, social cognition would have needed to take in proxy objects like signs, causal indexes, and symbols in the evaluation of intention. The explosion in mediated forms of communication since then has only amplified this dimension of social communication. In such a scenario, an intuition for deception is no good, because there is no stable vehicle (like facial expression) for it to attach to. Instead, one would once again need to explicitly deploy the machinery of Bayesian inference to arrive at an informed opinion concerning the social environment. No doubt, displacing this cognitively demanding task onto cultural tools like narrative solves some of the problem for the 4E theorist (Hutto 2011; 2009; 2008). But even so, on-the-fly social activities like strategizing, political maneuvering, and game-playing require formulating and evaluating beliefs using indirect evidence. Thus, while social cognition may begin with perceptually delivered and culturally supported intuitions about the mental states of others, these intuitions are simply not adequate to the complexity of social relations—for this, explicit theoretical modelling is also needed.

*The AI objection:* If theories are sometimes true without being useful, it can be because they are about phenomena that we cannot easily manipulate—this is the case with cosmology, for instance. When a theory seems to have no practical issue for a phenomenon that we can manipulate, it raises questions about the theory. With respect to 4E, this lack of practical impact is especially visible in relation to research in artificial intelligence. Though there were promising early interactions between the two (e.g. Brooks 1991; Clark 1998), the fact is that recent advances in areas like deep learning have been achieved without any reference at all to 4E perspectives. This is for the good reason that the architecture of most sophisticated deep learning models is based on reinforcement learning (Rescorla and Wagner 1972; Francois-Lavet et al. 2018)—an

avowedly computational paradigm that optimizes reward by minimizing some well-defined cost function. To be sure, the 4E theorist will object (and the AI researcher may agree) that organic cognition and artificial cognition are not the same thing, and there is no necessary reason that one need resemble the other. But then the stunning success of reinforcement learning theory in modelling both organic and artificial intelligence needs to be accounted for. As the computational commitments of reinforcement learning make mere adjustments to the theory ideologically unacceptable, this means reinforcement learning needs to be replaced with a 4E-friendly alternative. Friston, Daunizeau, and Kiebel (2009) recognize this problem and show that indirect optimization problems like the mountain car problem can, in principle, be solved using active inference models without the need for the computational apparatus of reinforcement learning or control theory. In a similar move, Clark (2016, 111–67) updates previous work on theoretical robotics from a predictive coding perspective with a view to avoiding cost functions. The fact remains, however, that replacing reinforcement learning models with active inference has only been demonstrated for relatively simple systems in environments that are fully specified in advance. There is (to my knowledge) no successful program of work in AI that systematically explores how non-reinforcement perspectives might be used to better design learning strategies for artificial agents in the computationally intractable environments where reinforcement learning has achieved its most striking successes (Silver and Hassabis 2017). Nor, for that matter, do any of the mainstream approaches in robotics seem to make lasting use of anything other than classical models of cognition (Lynch and Park 2017; Murphy 2019), despite the occasional vogue for approaches like Rodney Brooks’s “nouvelle AI.” Both AI and robotics are well-funded areas, and any perspective that promises competitive advantage is unlikely to remain unused for long. Thus, while the fact that 4E perspectives have not been found useful by AI researchers does not invalidate them, it does raise questions about their scope of application.

*The mathematical objection:* Though the rule is not absolute, the more mature a field of inquiry is, the more it lends itself to mathematical formalization. The value of this formalization is that it forces human intuition into contact with unanticipated conclusions. Think, for instance, of how the difficulty of proving Euclid’s parallel postulate eventually led to the discovery of the geometrical frameworks used in general relativity, or how the Dirac equation predicted the existence of the positron on formal grounds alone. It is unsurprising, therefore, that the most recent work in 4E cognition should reference the free energy principle (FEP), which allows for enactivist accounts of cognition to be mathematically formalized. The FEP is notoriously resistant to summary, but its core claim is that agents preserve their organizational integrity by minimizing information-theoretic free energy, which is the sum of two quantities: the negative log probability (termed “surprisal”) on encountering the range of sensory inputs the agent encounters, and the divergence between its model of the environment and the updated model of the environment given new data (Buckley et al. 2017). The mathematical relationship between these quantities—surprisal cannot exceed divergence—means that the FEP identifies, at least in theory, fundamental principles that seem to be present in most inference generating systems (Friston 2010; 2009). Where this connects with 4E is that, by mandating that the organism act on its environment so as to change sensory input



and thereby minimize model divergence, the FEP lodges cognition squarely in the body that enacts it.

But the question is, at what cost? For there is simply no way around the fact that the FEP brings back in all the computational and inferential machinery that 4E sought to obviate the need for in the first place. The FE-minimizing organism constructs a representation of states of the world over which it defines a probability distribution—the exact definition of a statistical model (McCullagh 2002). Equally, the quantification of model divergence using measures like cross entropy, precision, accuracy, and other metrics familiar from machine learning and information theory shift the understanding of cognition into a computational framework. Thus, the 4E theorist must pay for a formalized notion of embodied cognition by relinquishing the claim that cognition is not computational in nature. Some 4E theorists will be prepared to accept this trade-off; others will not. Either way, what remains clear is that 4E approaches face a dilemma: either eschew the apparatus of mathematical formalism and retreat into descriptive impotence or embrace it and inherit a computational model of cognition; *tertium non datur*.

No doubt, the determined 4E theorist will have answers to all of these objections—largely, because they are a matter of degree, rather than kind. This means that the issue becomes less one of arguing for a specific view of cognition, than one of arguing for where the emphasis of cognition lies. But this way grants all the *élan* of revolution with none of the drudge of responsible administration. And as Thomas Kuhn observes, the piling up of non-fatal objections can itself amount to a kind of fatal objection (Kuhn 1962). For this reason, if no other, it is easy to see why some critics take a harsh line with 4E. In one assessment, “for the vast majority of classic findings in cognitive science, embodied cognition offers no scientifically valuable insight” (Goldinger et al. 2016, 959). Another critic notes that “we have learned a tremendous amount about how perceptual systems process information by studying the brain, and, crucially, nothing seems to be left out” (Prinz 2006, 17). None of this is enough to damn an entire field of inquiry, but the skepticism is welcome in an area where enthusiasm all-too-frequently runs ahead of results.

The Oxford Handbook of 4E Cognition (OH4E) is a substantial tome, coming in at nearly 1,000 pages and forty-eight contributions. As a handbook, it is not designed to read through cover to cover—which is as well, because no one not in prison has that much time on their hands. The most useful innovation in this volume is the inclusion of critical reflections at the end of each section. These will prove especially valuable for the reader who is less familiar with the field, given that points of weakness in some essays are not especially obvious. (Nevertheless, more insistent editing would have eliminated the “I don’t have enough space to make my point” complaints that are sometimes visible in the reflections.)

The “What is Cognition?” section of the volume opens the author contributions with several big-picture reflections on the implications of 4E. Ezequiel Di Paolo, for instance, offers an intriguing reflection on how enactivism can be used to frame a philosophy of

life, while other contributors engage with the nature of extended cognition and the negotiation of multiple affordances. However, it is Dan Hutto and Erik Myin's chapter that is most provocative, in that it seeks to "abandon the information-processing and representationalist views of cognition in favor of a purely embodied know-how" (Hutto and Myin 2018, 105). That this is unlikely to succeed is made clear by their unconvincing handling of information-as-covariance. Despite this being the only successful model of representation in existence, the authors ask, "How can relations that hold between covarying states of affairs be literally 'extracted' and picked up from the environment so as to be 'encoded' within minds?" (2018, 101). If an answer were actually being sought here, the authors might do well to attend to research in deep learning, which explicitly designs artificial networks around information-extracting processes in the brain (Hassabis, Summerfield, and Botvinick 2017).

The "Modelling and Experimentation" section of OH4E provides what could have been a fruitful engagement with the highly contentious issue of model construction in 4E. Jakob Hohwy's chapter on predictive processing opens the section well enough with a candid concession of the tension between 4E and predictive coding models and an attempt to bridge the gap using Judea Pearl's notion of "Markov blankets." The issue is whether the deflated account of 4E that emerges will satisfy its most vigorous proponents. The subsequent contributions, however, do not do a great deal to advance thinking on either modelling or experimentation. The problem is stated baldly enough in Tom Froese's chapter, when he acknowledges that "there continues to be a lamentable disconnect between this growth in phenomenological and theoretical research on the one hand and the relative lack of experiments on genuine intersubjectivity on the other" (2018, 173). His response to this is to cite the perceptual crossing experiments of Auvray, Lenay, and Stewart (2009) and allude to some results from agent-based modelling. The intentions are good, but the perceptual crossing results are overworked (Di Paolo and De Jaegher also rely on them in EEC), and the degrees of freedom available for parameter selection in agent-based modelling too often makes it the last refuge of the scoundrel (I've authored enough to know).

Things start oddly in the "Cognition, Action, and Perception" section of the volume, in that we are once again presented with an essay (this one by Michael Kirchoff) that explores the relation between 4E and predictive coding. Why this should be needed in addition to Hohwy's earlier chapter is not clear, unless it's that the savage god of 4E must be placated with Kirchoff's more orthodox genuflections. Matthew Ratcliffe's essay on touch provides a judicious appreciation of a topic that is of fundamental importance to the 4E project, complements his previous work on depression. Joel Krueger also offers a worthwhile survey of direct social perception, usefully working through the various objections to the claim that we directly perceive emotions in others rather than indirectly infer them. But for reasons outlined above, this can only be achieved by shunting the inferential machinery somewhere else and asserting its irrelevance—whether it's conscious or not is beside the point.

Mark Rowlands opens the "Brain-Body-Environment Coupling and Basic Sensory Experiences" section with a confession: "There is a view of cognitive processes that I've

been peddling for more years than I care to remember (Rowlands 1995, 1999, 2002, 2003, 2006, 2010, 2011, 2013, 2015a, 2015b)" (Rowlands 2018, 350). As I was one of the undergraduates howled at by his wolf in 1990s Cork (Rowlands 2010b), this is a collective memory that dismays us both. Unfortunately, the OH4E essay does not do much to extend Rowlands's back catalogue into new areas, but there's no denying its value as an authoritative appreciation of what's needed to defend the 4E view of cognition. Like Rowlands's, Shaun Gallagher's chapter offers a worthwhile reprisal of his previous work, even if the solemn assurance that "the brain is part of the body, and has always been part of the body" (Gallagher 2018, 361) must count as one of the more inane trivialisms of 4E writing.

The essays in the "Social Cognition" and "Situated Affectivity" sections consist of eleven essays that take up the broad challenges posed by both topics. To me, the difference between the two sections does not clearly emerge from the chapters that comprise them, and the chapters themselves cover much well-trodden ground. These are two sections where the critical reflection section does particularly useful work in disentangling a nonobvious mélange of claims and counterclaims.

In the "Language and Learning" section, the emphasis falls on two topics that are challenging for the 4E paradigm. Perception and action are "embodied" in an obvious, hard-to-deny way; language and learning are not. Mark Johnson opens the section with an accessible summary of why language rewards being thought of as embodied, even if his case would have been even stronger had he made reference to the word norms literature showing how words are encoded using emotional and sensorimotor schemes (e.g. Brysbaert, Warriner, and Kuperman, 2014; Lynott and Connell, 2009; Warriner, Kuperman, and Brysbaert, 2013). Indeed, this latter criticism could be directed with even greater force at Michiel van Elk and Harold Bekkering's attempt to ground concept acquisition in predictive processing, which fails to mention the word norm data that provides exactly the evidence they need to ground their claims. (I should say that I've worked in word norm research, so I may be overly sensitive to the omission.) Of the other chapters, Marco Schmidt and Hans Rakoczy's essay on normativity in human behavior stands out as a nuanced appreciation of a challenging topic, even if—the occasional reference to shared intentionality notwithstanding—it doesn't seem to take from or add much to the 4E paradigm.

Louise Barrett opens the "Evolution and Culture" section with a defense of the basic-minds-without-content thesis previously advocated by Hutto and Myin. Barrett's wide range of ethological reference and more accessible exposition makes this seem a much more attractive option than when expressed by Hutto and Myin. However, the inevitable concession that "some of the interesting issues of human psychology are best understood using a representational frame-work" (2018, 730) duly arrives, inviting one to query whether "all" should be substituted for "some" to the disadvantage of the minds-without-content approach. The other highlight of the section is Kim Sterelny's chapter on how material culture impacted on cognition in deep time. Its value has less to do with any theoretical flashiness than its no-nonsense appreciation of how different classes of material and social artefact have shaped and amplified human thinking

across the paleo-anthropological record. If the contact with empirical evidence here is archaeological rather than experimental, it is no less welcome for that.

The final section of OH4E, “Applications,” is the least satisfying, if only because it seems like a dumping ground for essays that didn’t easily fit into the other sections. Topics range through psychopathology, intuitive rationality, robotics, philosophy of law, the humanities, and embodied aesthetics. Of these, Amy Cook’s piece on the humanities had the best claim on my interest. In it she does well to cover as much ground as she does, but even at that, the emphasis on literary and theatrical approaches might still be broadened to include other disciplines. Equally, the uncritical valorization of 4E approaches in the humanities might be improved by more critical reflection on the challenges that emerge from these applications—both with respect to how the scientific perspectives of 4E might be legitimately (or illegitimately) used in the humanities, and how the humanities might broaden the perspectives of 4E (and cognitive science generally). This is especially needful now that digital and experimental methods are, for the first time, making interpretive hypotheses testable. The humanities are in flux as they have never been before, and if they are to have a future (which one hopes they do) it cannot be one in which 4E just replaces, say, Marxism or psychoanalysis as an interpretive framework—methodological innovation is also needed (Carney and Troscianko 2014). OH4E ends on a pleasant note with an informed and lively investigation by Barbara Gail Montero of how proprioception might inform aesthetic judgment.

So where now for 4E cognition? To start with, it is no longer the future and probably needs to stop saying so. The shambling, wheezing Brezhnev awarding himself the Order of Lenin for zeal in the cause of revolution is not a good look, but that is what 4E resembles when it burnishes its revolutionary credentials thirty years after the revolution has ended. Secondly, the essays I have surveyed here make it quite clear that, even if the claims of 4E can in principle be empirically supported, they are not yet sufficiently supported by the evidence available. Theoretical speculation is obviously valuable, but at some point, empirical claims about the world (and the minds that it accommodates) need to be measured against that world—and no amount of phenomenological reflection is up to compensating for that. And finally, 4E needs to ditch the covert theology. As Jesse Prinz notes, “scientists have had to work hard over many, many years to support the conclusion that experiences are in the head” (2006, 16). The subjective attractiveness of taking them back out again should make us suspicious, and this means being especially careful that 4E approaches are appealing to our reason and not our sensibilities. If this vigilance succeeds, 4E will merge with the wider body of hard-won truths about the human cognitive endowment. But if it fails—and failure is the norm—4E will join psychoanalysis, structuralism, functionalism, cybernetics and all those other perspectives on the mind that monopolized the correct answers right up to the point where they didn’t. This would be a disappointing fate for a paradigm that, for all its failings, still has a great deal to offer the study of cognition.

# The Power of the Powerless by Václav Havel

Václav Havel's *The Power of the Powerless* is fascinating and surprising. The purpose of this review is to summarize it in such a tantalizing (and lengthy) way that you stop reading this and go look up Havel's original instead.

Its context is Czechoslovakia in 1978. The country has been behind the Iron Curtain for thirty years, and ten years have passed since the brief experiment in political liberalization known as the "[Prague Spring](#)" which had been quickly choked off by a Soviet-led invasion.

[Havel](#) was a Czech playwright with international renown, whose works were banned in his own country. In 1977 he helped to spearhead "[Charter 77](#)" — a document that called on the government to respect human rights and its own Constitution.

Charter 77 was prompted by the prosecution of members of the rock band "[The Plastic People of the Universe](#)" — the "[Pussy Riot](#)" of their day. The government took the threat represented by the Charter seriously — it persecuted its signers and made it illegal to print or distribute the text. In 1979, Havel would be sent to prison for four years for his role in advocating for the Charter.

While Havel wrote *The Power of the Powerless*, he was under constant police surveillance and harassment for his Charter 77 activism. Meanwhile he was being noticed by freedom-loving people around the world and being hailed as a prominent Soviet-bloc dissident.

## The "Post-Totalitarian" System

He begins his argument by examining the system the "dissidents" are up against. It's a strange new form of tyranny — a dictatorship not by a person or people, but by a bureaucracy and by certain principles and external contingencies (the primary one being that the Soviet Union intends to maintain Czechoslovakia as an obedient client state). So this is not the sort of dictatorship that can be threatened by attacks on a particular person or clique.

Havel refers to this system as "post-totalitarian". This is confusing, as it seems at first to imply that it is no longer totalitarian, which isn't the case. "Neototalitarian" might have been more apt.

Czechoslovakia (and the Soviet Union and its client states collectively) was ruled in part by an ideology — almost a religion — that had proven to be a tempting refuge for confused, uprooted, and alienated people. "Of course, one pays dearly for this low-rent home: the price is abdication of one's own reason, conscience, and responsibility, for an essential aspect of this ideology is the consignment of reason and conscience to a higher authority. The principle involved here is that the center of power is identical with the center of truth."

He anticipates the argument that though this ideology is dominant and everpresent, few people *really* believe its platitudes. They're like schoolchildren reciting the Pledge of Allegiance without paying attention to the words because it's what is expected of them. To meet this argument Havel introduces us to a grocer:

# The Obedient Grocer

In the window of the grocery is a sign that reads “Workers of the world, unite!”

What does the grocer mean by putting this sign in the window? Not that he is enthusiastic about global worker unity and wants to spread the word about it. The *real* message on the sign reads something like this: “I, the grocer X—, live here and I know what I must do. I behave in the manner expected of me. I can be depended upon and am beyond reproach. I am obedient and therefore I have the right to be left in peace.” The message is not meant for the grocer’s customers, but for officials who might suspect him or for informers who might care to turn him in.

If the grocer had to put that implicit message *explicitly* in his window, he might be embarrassed to be seen kowtowing in such a way, but by genuflecting in this indirect manner he saves face. If you ask him why he has the sign in his window, he can answer “What’s wrong with the workers of the world uniting?” and protect his dignity. In this way a gesture of obedience and subservience is disguised by ideology as one of solidarity and empowerment.

## Ideology

Ideology, Havel summarizes, “offers human beings the illusion of an identity, of dignity, and of morality while making it easier for them to part with them.... It is a veil behind which human beings can hide their own fallen existence, their trivialization, and their adaptation to the status quo. It is an excuse that everyone can use, from the greengrocer, who conceals his fear of losing his job behind an alleged interest in the unification of the workers of the world, to the highest functionary, whose interest in staying in power can be cloaked in phrases about service to the working class. The primary excusatory function of ideology, therefore, is to provide people, both as victims and pillars of the post-totalitarian system, with the illusion that the system is in harmony with the human order and the order of the universe.”

Today’s tyrannies are too large and complex and cannot be held together by force and fear alone. They require their subjects not merely to submit passively to but to participate actively in their own subjection. Ideology is the mechanism to accomplish this.

Individuals need not believe all these mystifications, but they must behave as though they did, or they must at least tolerate them in silence, or get along well with those who work with them. For this reason, however, they must live within a lie. They need not accept the lie. It is enough for them to have accepted their life with it and in it. For by this very fact, individuals confirm the system, fulfill the system, make the system, are the system.

Once you make the decision to adopt the ideological mask for your subservient behavior — like the grocer putting the sign in his window — you become a part of this glue that affixes ideology over reality and gives ideology power. It doesn’t matter that you inwardly don’t really believe the explicit message of the ideology, because the *explicit* message isn’t the important one, and it doesn’t matter if you believe it or not so long as you agree to continue acting as though you did.

While ideology is central to the post-totalitarian power structure, the interests of the structure itself are paramount, and the ideology — or the interpretation of it anyway — will tend to be subordinate to it. The tighter the control that the government exercises over communication and expression, the better it will be able to enforce and manipulate the orthodox interpretation of the

ideology and the more the ideology will come to float far above reality, more-or-less completely detached from it: “a world of appearances, a mere ritual, a formalized language deprived of semantic contact with reality and transformed into a system of ritual signs that replace reality with pseudo-reality.”

For example, in China today, “communism” is still the name given to the sacred ideology that is said to govern the system, but its meaning has come a long way: now it means the total state-enforced subjection of the working class to a small minority of fantastically wealthy private owners of the means of production (the wealth of China’s National Congress makes the U.S. Congress look like a bunch of ordinary middle-class schmoes). It’s still “communism” you’re expected to be loyal to, the flag is still red, that’s still Mao’s face staring back at you from the money, and you can still signal your loyalty to the system with the same empty platitudes about the rule of the working class — but the system doesn’t care about the explicit meaning of the platitudes any more than you do.

Because ideology can become so absurdly detached from reality in this way, it can be a real art to try to maintain your fiction of adherence to it. Because of this, “the virtuosity of the ritual” comes to be more important than actually being able to attach meaning to what you are doing or saying. Aspects of the ritual and of the ideology come to represent only each other and never come down to earth. This can cause the ideology to detach even from the bureaucracy it was designed to serve, until it becomes an independent, malignant, power-appropriating memetic menace all its own.

At this stage, when the ideology is serving itself more than it serves the bureaucracy, the power structure stops attracting the ambitious and starts attracting the faceless — empty suits — people who can articulately and cleverly engage in the virtuosity of the ritual but who don’t seem to have much going on outside of this arena of rhetorical swamp gas and who are so thoughtless that they have thoroughly internalized the ideology’s criteria of success and prestige.

If a Frankenstein’s Monster ideology like this is so powerful that it can eventually even conquer and press into service the post-totalitarian dictatorship itself, what hope do we have? It is this: the ideology “is built on lies [and] works only as long as people are willing to live within the lie.”

## Who Enforces the Ideology?

What if our grocer were to stop living in the lie in one little way: by not hanging the sign in his window that means nothing to him. Well, what possible difference could that make? It’s unlikely any of his customers even notice the sign. The sign is not meant to be read individually, anyway, but “to form part of the panorama of everyday life.” It is as a contributor to this panorama that the grocer serves the system. The message of the panorama is not the message on the sign but this message: “this sign-hanging is what the ideology demands of us today and we are complying.” Those who hang the signs are complying with the ideology and expressing the ideology’s demands by the same action: simultaneously the voice of command and the posture of submission.

This has a pernicious psychological effect. The latent consciousness that you are both victim and perpetrator of this ideological control influences you to *identify* with the ideology. You feel better both in submitting and in commanding if you think you are doing so in service of an ideology you believe in, so you try to believe that you believe in this weird, untethered, nonsensical ideology — and you come to see attacks against the ideology as threats to you personally.

Thus the conflict between the aims of life and the aims of the system is not a conflict between... the rulers and the ruled.... In the post-totalitarian system, this line runs *de facto* through each person, for everyone in his own way is both a victim and a supporter of the system. What we understand by the system is not, therefore, a social order imposed by one group upon another, but rather something which permeates the entire society and is a factor in shaping it, something which may seem impossible to grasp or define (for it is in the nature of a mere principle), but which is expressed by the entire society as an important feature of its life.

...It can happen and did happen only because there is obviously in modern humanity a certain tendency toward the creation, or at least the toleration, of such a system. There is obviously something in human beings which responds to this system, something they reflect and accommodate, something within them which paralyzes every effort of their better selves to revolt. Human beings are compelled to live within a lie, but they can be compelled to do so only because they are in fact capable of living in this way. Therefore not only does the system alienate humanity, but at the same time alienated humanity supports this system as its own involuntary master plan, as a degenerate image of its own degeneration, as a record of people's own failure as individuals.

Alongside the striving for dignity, integrity, and personality that we naturally value and treasure in ourselves, says Havel, there lives a less-acknowledged, sinister striving "to merge with the anonymous crowd and to flow comfortably along with it down the river of pseudolife."

## The Disobedient Grocer

So if going along with the lie means not just *submitting* to the system, but *enforcing* the system, what is the alternative for people who find the system intolerable? What if the grocer stops participating in the lie and starts living in the truth?

Well, first and most obviously, the system retaliates against him. And since the system is enforced not just by the officials who overtly persecute him but by everyone, they as part of their life-in-the-lie must also shun him. Does he then just become a vivid display of the danger of rebellion — useful to the regime and no threat to it? According to Havel, no: the grocer's quixotic act is indeed a serious threat to the system:

By breaking the rules of the game, he has disrupted the game as such. He has exposed it as a mere game. He has shattered the world of appearances, the fundamental pillar of the system. He has upset the power structure by tearing apart what holds it together. He has demonstrated that living a lie is living a lie. He has



broken through the exalted façade of the system and exposed the real, base foundations of power. He has said that the emperor is naked. And because the emperor is in fact naked, something extremely dangerous has happened: by his action, the greengrocer has addressed the world. He has enabled everyone to peer behind the curtain. He has shown everyone that it is possible to live within the truth. Living within the lie can constitute the system only if it is universal. The principle must embrace and permeate everything. There are no terms whatsoever on which it can co-exist with living within the truth, and therefore everyone who steps out of line denies it in principle and threatens it in its entirety.

The biggest vulnerability of the ideologically-ruled post-totalitarian system is its unmooring from reality. But this vulnerability only becomes a liability when the system is brought into contrast with reality — and the system works hard to ensure that this doesn't happen; that's the point of the panorama of platitudes and of the conscription of the grocer to help deploy it.

In the post-totalitarian system, therefore, living within the truth has more than a mere existential dimension (returning humanity to its inherent nature), or a noetic dimension (revealing reality as it is), or a moral dimension (setting an example for others). It also has an unambiguous political dimension. If the main pillar of the system is living a lie, then it is not surprising that the fundamental threat to it is living the truth. This is why it must be suppressed more severely than anything else.

## Truth-Force

Havel believes the lie can never gain total control of someone. There's always a seed of truth remaining. The lie, because it so vigorously claims to be the truth and claims to be a route to authenticity, reinforces the idea that there *is* such a thing as truth and authenticity and that these things are valuable. The truth, then, is not exterminated in the post-totalitarian system, but continues to flow through society like an underground stream. Those who decide to live in the truth are not, therefore, isolated and having to invent themselves from scratch, but they're able to tap into this reservoir.

Because the post-totalitarian system is so detached from and hostile to the truth, the flow of this underground stream can evade its notice, "and by the time it finally surfaces into the light of day as an assortment of shocking surprises to the system, it is usually too late to cover them up in the usual fashion. Thus they create a situation in which the regime is confounded, invariably causing panic..."

This manner of striking at the main vulnerability of the post-totalitarian ideological system is a peculiar form of opposition: it doesn't take place in the halls of power or in the voting booth or in conspiratorial revolutionary cells or in strikes and street protests, but at "the level of human consciousness and conscience, the existential level... in the fifth column of social consciousness, in the hidden aims of life, in human beings' repressed longing for dignity and fundamental rights... This power does not participate in any direct struggle for power; rather, it makes its influence felt in the obscure arena of being itself." But once established there, it can and does contribute, in subtle but definite ways, to such things as "a social movement, a sudden explosion of civil unrest, a sharp conflict inside an apparently monolithic power structure, or simply an irrepressible transformation in the social and intellectual climate" and thereby has powerful political consequences.

The Prague Spring itself, Havel says, only superficially was a conflict between groups vying for political power. Looked at more closely, it appears as “the final act and the inevitable consequence of a long drama originally played out chiefly in the theatre of the spirit and the conscience of society” and prompted by a few individuals with no pretensions to political power who simply decided to begin living in the truth. The Prague Spring wasn’t the birth of something promising that was then cut down, but the above-ground blooming of something that continues to spread its mycelia underground.

And this is why these post-totalitarian ideological systems are so intolerant of leaks and dissent. Why was Solzhenitsyn hounded out of Russia? For the same reason Ed Snowden was hounded into it: “a desperate attempt to plug up the dreadful wellspring of truth, a truth which might cause incalculable transformations in social consciousness, which in turn might one day produce political debacles unpredictable in their consequences.”

But living the truth isn’t just a matter of exposing facts. “It can be any means by which a person or a group revolts against manipulation: anything from a letter by intellectuals to a workers’ strike, from a rock concert to a student demonstration, from refusing to vote in the farcical elections to making an open speech at some official congress, or even a hunger strike, for instance.... every free expression of life indirectly threatens the post-totalitarian system politically, including forms of expression to which, in other social systems, no one would attribute any potential political significance, not to mention explosive power.” There’s a reason why Charter 77 was prompted by the prosecution of a rock-and-roll band.

## Living in the Truth as a Moral, not Political Revolution

Havel thinks the political scenario he has described is a symptom of a widespread epidemic of ethical pathology. The fact that living-in-a-lie has emerged as a self-perpetuating political system suggests that something has gone badly wrong at the core of society and in the moral centers of the people who make it up. It reveals that we have demoralized ourselves by abandoning our senses of responsibility in order to dissolve our identities in the solvent of mass culture and consumerism.

Seen in this light, the political side effects of living in the truth are secondary to its function of allowing us to reclaim our moral agency. Indeed, because the beneficial political effects of living in the truth are so diffuse and difficult to trace, and the consequences of confronting the system in this way are in contrast so personal and visceral and likely, nobody would be likely to make the attempt if there were not this additional imperative.

## Politics Under Post-Totalitarianism

How can a citizen participate in the political process that governs his or her society? In the post-totalitarian system, the normal way to participate is by living in and helping to enforce the lie that perpetuates the stranglehold of the system over the lives of its subjects. Parts of this lie are the farcical processes — voting or pleading with representatives and so forth — that mimic (and, according to the lie, constitute) political deliberation and action. The only form of participation permitted to you is one that helps you propel the system that smothers politics, not one that actually allows you to make decisions together with your fellow citizens.

What if you want more than that: participatory politics of equals, rather than the obedient pseudopolitics of the galley slave? Do you participate in the fake elections more vigorously?

lobby your fake representatives more persuasively? These things are hopeless and dangerous and make people cynical: if you don't see beyond the officially-sanctioned outlets of pseudopolitics, there seems to be no point to politics at all.

Living in the truth is the remaining method of political activity — the last alternative to the pseudopolitics that just makes things worse. This can sometimes trip up well-meaning activists, who may overestimate the usefulness of confrontational and bold “political” acts — that is, acts within the permitted pseudopolitical sphere — and thereby bolster the perceived legitimacy of that sphere. This is another way of participating in the lie. Instead, effective activists need to understand that they're in a new sort of system with new rules, and they need to be imaginative and not try to build within either the pseudopolitical framework or within old-fashioned models of dissent.

To foment an opposition, don't paint a picture of a better set of rulers or a new political party or a constitutional amendment or electoral reform or any of that perennial hogwash. Instead, aim concretely and directly at “the continuing and cruel tension between the complex demands of [the post-totalitarian] system and the aims of life, that is, the elementary need of human beings to live, to a certain extent at least, in harmony with themselves... in a bearable way, not to be humiliated by their superiors and officials, not to be continually watched by the police, to be able to express themselves freely, to find an outlet for their creativity, to enjoy legal security, and so on.”

We cannot free ourselves by overthrowing a tyranny and imposing freedom from above the way communism was imposed on us from above; instead we have to strive to become free and then impose our freedom on the government from below.

## Dissent and Opposition

In a democracy, the opposition is a party currently out of power working through legitimate channels within the system to try to gain or exert power. In a traditional dictatorship, the opposition is those people who are trying to replace the dictatorship with something else.

Havel says that Charter 77 is not an opposition in these senses, though some of the signers may have aspirations in this direction. It isn't a political party with aspirations of gaining political power, and it doesn't have an alternative system it hopes to install in place of the present state. Nonetheless, “Western journalists” have seized on Charter 77 as an “opposition movement” and the Czech government treats it as an oppositional organization simply because it “manages to avoid total manipulation and... therefore denies the principle that the system has an absolute claim on the individual.”

“Opposition” is a tricky word. Once the label gets attached to you, you can expect hassles from the state: you are considered a traitor and can expect treatment ranging from character assassination to outright execution. But it's also deceptive in that it defines your work not in its own terms or in how it relates to reality, but in terms of the system of lies you're trying to escape: rather than living in the truth you find yourself defined as living in opposition to the lie, and in a way that becomes contaminated with the lie (sources say these protesters oppose the workers of the world uniting).

Some people who are trying to live in the truth in the Soviet bloc were called “dissidents” — something Havel belittles as a sort of Western media-granted celebrity status (he always puts the word in quotation marks).

One danger of this label is that it comes to sound like a profession — like you have to have a license to dissent, or like it's only for people who have made it their special vocation. In fact, “dissidents” are not people who “consciously decided to be professional malcontents” but ordinary people “who are doing what they feel they must and, consequently, who find themselves in open conflict with the regime.”

The label has a way of separating a small group of people into an elevated clique and treating them like a tiny interest group distinct from society at large: journalists ask “is the government going to respect the rights of the dissidents” rather than “is the government going to respect everyone's right to live in the truth?”

It is truly a cruel paradox that the more some citizens stand up in defense of other citizens, the more they are labeled with a word that in effect separates them from those “other citizens.”

## Small-Scale Work

What of the argument that it's worth making small concessions to the lie in order to be granted the limited freedom and resources necessary to do good work? Why not work within the system and try to make it better or to ameliorate its problems?

There's something to this: “It is hard to say how much worse things would be if there were not many hard-working people who simply refuse to give up and try constantly to do the best they can, paying an unavoidable minimum to living within a lie so that they might give their utmost to the authentic needs of society. These people assume, correctly, that every piece of good work is an indirect criticism of bad politics, and that there are situations where it is worthwhile going this route, even though it means surrendering one's natural right to make direct criticisms.”

But Havel says that this option has become less tenable in Czechoslovakia. Things have become too rotten. The compromises are too overwhelming. Too much good work ends up being hijacked and parasitized to feed the corrupting engine of the system. When the system requires total adherence to an ideology that has become totally unmoored from the truth, how much good can you do without butting up against the ideology's limits? If you decide to stay safe, you lose your ability to do good; if you decide to keep doing good, you find yourself suddenly a “dissident” in spite of your modest intentions.

But this is not one-size-fits-all advice. If you find that in your situation you can do the most good by making tactical concessions to the lie, Havel advises that you make your judgment call and do what you can. It is possible to live honorably this way. If you do the right thing and find out that (surprise!) it's also permitted — that's a marvelous discovery.

Living in the truth is not necessarily overtly oppositional or dissenting at all. Some of it is subtle and not particularly visible — “you simply straighten your backbone and live in greater dignity as an individual.” Other parts are more visible and shared: “everything from self-education and thinking about the world, through free creative activity and its communication to others, to the most varied free, civic attitudes, including instances of independent social self-organization.” When there is enough of this going on, it forms the soil in which more overtly and consciously political initiatives can grow.

A movement of “dissent” requires a healthy substrate of independent, grassroots social activity and organization; this in turn depends on individuals willing to lead such independent ways of living by living in the truth as individuals even in the absence of this social support structure.

But remember those quotation marks around “dissent” — it’s not so much that self-consciously dissident groups are going to emerge from this strata of independent ways of living, but that some of these independent ways of living are going to be persecuted by an intolerant government and will thereby become dissident activities.

## People and Politics

The “dissident” movements in the Soviet bloc, Havel says, are defensive: that is, they are defending human beings against a smothering anti-human system. He contrasts this with political movements, which may have an offensive as well as defensive program, for instance a program to institute a different sort of system or to reform the existing one in a particular way.

Havel thinks this is not a liability but an advantage: “it forces politics to return to its only proper starting point... individual people.” He thinks that things have gotten so bad in his country that the central issue isn’t about what shape the political system ought to take but about what to do for the people who are victimized by the political system. In contrast: “In the democratic societies, where the violence done to human beings is not nearly so obvious and cruel, this fundamental revolution in politics has yet to happen, and some things will probably have to get worse there before the urgent need for that revolution is reflected in politics.”

Every society, of course, requires some degree of organization. Yet if that organization is to serve people, and not the other way around, then people will have to be liberated and space created so that they may organize themselves in meaningful ways. The depravity of the opposite approach, in which people are first organized in one way or another (by someone who always knows best “what the people need”) so they may then allegedly be liberated, is something we have known on our own skins only too well.

## Revolt and Law

Maybe it’s time to revolt, to overthrow the system entirely and to install a new one by force. But such a revolt is difficult to imagine in the post-totalitarian system. Such a revolt would involve two opposed forces of roughly equivalent strength meeting in the arena of actual force and political power. But in the post-totalitarian system:

Society is not sharply polarized on the level of actual political power, but, as we have seen, the fundamental lines of conflict run right through each person. In this situation, no attempt at revolt could ever hope to set up even a minimum of resonance in the rest of society, because that society is soporific, submerged in a consumer rat race and wholly involved in the post-totalitarian system (that is, participating in it and acting as agents of its automatism), and it would simply find anything like revolt unacceptable. It would interpret the revolt as an attack upon itself and, rather than supporting the revolt, it would very probably react by intensifying its bias toward the system, since, in its view, the system can at least guarantee a certain quasi-legality. Add to this the fact that the post-totalitarian system has at its disposal a complex mechanism of direct and indirect surveillance

that has no equal in history and it is clear that not only would any attempt to revolt come to a dead end politically, but it would also be almost technically impossible to carry off. Most probably it would be liquidated before it had a chance to translate its intentions into action. Even if revolt were possible, however, it would remain the solitary gesture of a few isolated individuals and they would be opposed not only by a gigantic apparatus of national (and supranational) power, but also by the very society in whose name they were mounting their revolt in the first place. (This, by the way, is another reason why the regime and its propaganda have been ascribing terroristic aims to the “dissident” movements and accusing them of illegal and conspiratorial methods.)

“Dissident” movements tend to have a strong bias against violent change, though not one that veers dogmatically into pacifism:

“[D]issidents” tend to be skeptical about political thought based on the faith that profound social changes can only be achieved by bringing about (regardless of the method) changes in the system or in the government, and the belief that such changes — because they are considered “fundamental” — justify the sacrifice of “less fundamental” things, in other words, human lives.

The “dissident” view is that you don’t change the system first, but that the system will change incidentally as changes take place in the people who uphold and evolve the system. They “do not shy away from the idea of violent political overthrow because the idea seems too radical, but on the contrary, because it does not seem radical enough.”

Thus an attitude that turns away from abstract political visions of the future toward concrete human beings and ways of defending them effectively in the here and now is quite naturally accompanied by an intensified antipathy to all forms of violence carried out in the name of a better future, and by a profound belief that a future secured by violence might actually be worse than what exists now; in other words, the future would be fatally stigmatized by the very means used to secure it.

Havel notes that many of the “dissident” groups, like the Charter 77 movement, claim to be acting in the defense of various doctrines of international or national law — “such as the Universal Declaration of Human Rights, the International Covenants on Human Rights, the Concluding Act of the Helsinki Agreement, and the constitutions of individual states.” Why might this be? After all, it is completely naïve to think that their governments actually respect these laws, have any interest in respecting them, or can be compelled by some greater force to respect them. Is pretending that the law is meaningful just another way of living in the lie?

Havel defends this legalistic approach, at least as it applies to the post-totalitarian system.

Such a system is not dominated by power-wielding groups or individuals, the way a traditional totalitarian system is, but by bureaucracy and ideology. It “is utterly obsessed with the need to bind everything in a single order: life in such a state is thoroughly permeated by a dense network of regulations, proclamations, directives, norms, orders, and rules.” The legal code is one expression of the ideology underlying the system, and of one of the lies of the system: that it is well-regulated, governed by law, eager for justice, and vigorous in its defense of human rights.

If an outside observer who knew nothing at all about life in Czechoslovakia were to study only its laws, he would be utterly incapable of understanding what we were complaining about. The hidden political manipulation of the courts and of public prosecutors, the limitations placed on lawyers' ability to defend their clients, the closed nature, *de facto*, of trials, the arbitrary actions of the security forces, their position of authority over the judiciary, the absurdly broad application of several deliberately vague sections of that code, and of course the state's utter disregard for the positive sections of that code (the rights of citizens): all of this would remain hidden from our outside observer....

But that is not all: if our observer had the opportunity to study the formal side of the policing and judicial procedures and practices, how they look "on paper," he would discover that for the most part the common rules of criminal procedure are observed: charges are laid within the prescribed period following arrest, and it is the same with detention orders. Indictments are properly delivered, the accused has a lawyer, and so on. In other words, everyone has an excuse: they have all observed the law.

The legal code is also the mechanism by which the various parts of the system communicate with each other and establish their roles and duties. It's a sort of scaffolding. "It provides their whole game with its rules and engineers with their technology.... Without the legal code functioning as a ritually cohesive force, the post-totalitarian system could not exist."

So this is the reasoning behind the legalistic approach. There's no need to pretend that the law is anything but what it is, but that doesn't mean that the law cannot be used to advantage. The system depends on it and, to some extent anyway, must flow through the channels it defines in order to function.

I have frequently witnessed policemen, prosecutors, or judges — if they were dealing with an experienced Chartist [(Charter-signer)] or a courageous lawyer, and if they were exposed to public attention (as individuals with a name, no longer protected by the anonymity of the apparatus) — suddenly and anxiously begin to take particular care that no cracks appear in the ritual. This does not alter the fact that a despotic power is hiding behind that ritual, but the very existence of the officials' anxiety necessarily regulates, limits, and slows down the operation of that despotism.

## Building Parallel Structures

There's another choice besides revolt and legalism. Rather than try to overthrow the current system or turn its rulebook against it, you can extend your participation in ways of life that substitute for the system's.

In a way this naturally follows from the independent ways of living mentioned earlier. As more people live in the truth and develop these independent ways of living, they will come to do so *together*, interacting and creating new ways of organizing and structuring these independent

activities. New organizations and structures will fill spaces that the State has left unfilled or that it fails to monopolize. Some examples of this in Czechoslovakia were the underground music scene and the samizdat publishing and distribution industry, but the form had potential to extend further, into such things as “parallel forms of education (private universities), parallel trade unions, parallel foreign contacts, to a kind of hypothesis on a parallel economy” and eventually a parallel state (Havel attributes [these ideas](#) to [Václav Benda](#)).

This approach is people-centered, it's not just aimed at some future benefit but is beneficial in the here-and-now, it's practical and not just theoretical, it's something everyone can participate in, and it's radical in the sense that it works directly at the root of people's day-to-day lives rather than in the superstructure of the system.

Because people who decide to live in the truth are, at first anyway, isolated rebels distinct from society, there is a temptation to see them as individualists in *retreat* from society: outcast or in isolation (the title “dissident” is another way of emphasizing this point of view). It is more accurate to see them not as retreating from society but advancing before it: as experimental pathfinders beating new trails and inviting society to follow their lead. Similarly, when groups of people develop parallel structures that substitute for those sanctioned by the system, this is not an act of monastic retreat or ghettoization but one of experimental advance.

So be careful not to see this parallel world as an end in itself, as though once we get it established we will be able to migrate there and leave the other world behind. So long as the post-totalitarian system rules, our participation in the parallel world will be tainted by the same schizophrenia that everyone suffers: trying to live with partial respect for the truth and partial subservience to the lie. The point is not to establish an underground in which you can enjoy furtive moments of freedom, but to free everyone (a sort of [“mahāyāna” agorism](#) perhaps).

The system will react to these parallel structures in two ways: by trying to repress them and by trying to coopt them. The repression is straightforward: practices will be banned, practitioners persecuted. Cooptation is a little more subtle. The system may adopt those aspects of the parallel world that are especially popular or effective or difficult to control. This may be a positive thing, something of a real reform, but it often is just a way of rendering the parallel world safe to the system — defanging it, integrating it into the lie, and slapping a patina of progress and liberality onto the system's façade. It can be confusing, if not deliberately baffling.

But the cooptation works both ways, and it will be through such grudging attempts at compromise that the post-totalitarian system will eventually be defeated. It is not the case either that all of these attempts at cooptation are bad and should be resisted or that all of them are good and should be encouraged. Is your attempt at a parallel structure partially contaminated by the ruling system? Of course it is. Don't let this discourage you; don't let the perfect be the enemy of the good. Keep your eyes on the truth and keep moving forward in the direction it points.

The post-totalitarian system and the parallel system of people developing ways of living in the truth together are two incompatible worlds, and one of them must go: “either the post-totalitarian system will go on developing... thus inevitably coming closer to some dreadful Orwellian vision of a world of absolute manipulation, while all the more articulate expressions of living within the truth are definitely snuffed out; or the independent life of society (the parallel polis), including the ‘dissident’ movements, will slowly but surely become a social phenomenon of growing importance, taking a real part in the life of society with increasing clarity and influencing the general situation.”



What will prompt the *coup de grâce* is impossible to predict. It will probably be some accident of history or the culmination of trends that are only clear in retrospect. Our task is not to plot this revolution but to lay the groundwork that makes it possible or inevitable.

## The Crisis of Contemporary Technological Society

The problem we're faced with is deeper than the specific post-totalitarian system in the Soviet bloc, and will not end when it does:

Technology — that child of modern science, which in turn is a child of modern metaphysics — is out of humanity's control, has ceased to serve us, has enslaved us and compelled us to participate in the preparation of our own destruction. And humanity can find no way out: we have no idea and no faith, and even less do we have a political conception to help us bring things back under human control. We look on helplessly as that coldly functioning machine we have created inevitably engulfs us, tearing us away from our natural affiliations (for instance, from our habitat in the widest sense of that word, including our habitat in the biosphere) just as it removes us from the experience of Being and casts us into the world of "existences."

Here, too, we need a revolution, but it must be more fundamental, not "merely philosophical, merely social, merely technological, or even merely political" but existential, "a generally ethical — and, of course, ultimately a political — reconstitution of society."

The post-totalitarian system is only one aspect — a particularly drastic aspect and thus all the more revealing of its real origins — of this general inability of modern humanity to be the master of its own situation. The automatism of the post-totalitarian system is merely an extreme version of the global automatism of technological civilization. The human failure that it mirrors is only one variant of the general failure of modern humanity.

Western liberal democracy — that famous ["end of history"](#) — is not an adequate response to this crisis. "It may even be said that the more room there is in the Western democracies (compared to our world) for the genuine aims of life, the better the crisis is hidden from people and the more deeply do they become immersed in it...."

People are manipulated in ways that are infinitely more subtle and refined than the brutal methods used in the post-totalitarian societies. But this static complex of rigid, conceptually sloppy, and politically pragmatic mass political parties run by professional apparatuses and releasing the citizen from all forms of concrete and personal responsibility; and those complex focuses of capital accumulation engaged in secret manipulations and expansion; the omnipresent dictatorship of consumption, production, advertising, commerce, consumer culture, and all that flood of information: all of it, so often analyzed and described, can only with great difficulty be imagined as the source of humanity's rediscovery of itself.... In a democracy, human beings may enjoy many personal freedoms and securities that are unknown to us, but in the end they do them no good, for they too are ultimately victims of the same automatism, and are incapable of defending their concerns about their own identity or preventing their superficialization or transcending concerns about their own personal survival to become proud and responsible members of the polis, making a genuine contribution to the creation of its destiny.

For this reason, it would be short-sighted for post-totalitarian dissidents to set their sights on establishing a democracy of this sort as anything but a temporary stepping stone to a society of dignity. Confronted with a “post-totalitarian” situation, we need a “post-democratic” solution.

## What Is to Be Done?

What we really need is a reconstitution of the larger human order of which the political order is just a part. And this means a society-wide moral revolution of such things as “[a] new experience of being, a renewed rootedness in the universe, a newly grasped sense of higher responsibility, a newfound inner relationship to other people and to the human community.... In other words, the issue is the rehabilitation of values like trust, openness, responsibility, solidarity, love.”

The political reformation will follow naturally from this. It is hazardous to try to predict in advance what it will look like, but there are some aspects that we might anticipate: It will probably rely more on smaller units of organization that are based on natural communities of people with shared interests (rather than big states with arbitrary geographical boundaries). These units of organization will not have monopolistic impulses but will be welcoming of new and of parallel structures. They will be less formal — not like organizations but like communities. Their authority will be based on their utility, not on sovereignty. They will be more likely to spring up ad-hoc as needed rather than being on-going institutions. Individuals who have authority will not have it by virtue of their title or position, but because of the trust they have earned in taking responsibility for the specific tasks at issue. This may mean that they have *more* political power than the politicians today.

And this goes for economic organization as well as political organization: Self-managed, purposeful units — not autonomous, self-interested corporations with subservient workers who have no stake in or responsibility for their work.

And come to think of it... these little clusters of people living in the truth, finding each other, forming human bonds of solidarity by desperate necessity, creating experiments in parallel structures based on concrete human needs — aren't they demonstrations of the sort of world we're trying to feel our way toward?

# The Recovering by Leslie Jamison

In late 2018, I read *The Recovering* by Leslie Jamison, a compelling memoir of debasement and redemption. This book changed the way I see the world through the deep lessons it offers about how to tell powerful stories, and thus about how to build new culture of the kind we need in order to re-infuse meaning into modern life.

Jamison's voracious alcoholism devours the first decade of her adulthood, leaving her bereft and hopeless. From rock bottom, she slowly finds a path forward through AA, and over the subsequent years she takes her first halting steps towards (and away from, and back towards) recovery. It's beautifully written, sometimes salacious, and often heartrending.

Jamison learns that AA culture builds and maintains a collection of ancient narrative templates. Moreover, she discovers that when she maps the unique landscape of her own suffering onto those pre-built forms, sharing her story in AA groups with no goal in mind but the process of doing so, she is finally granted a form of grace powerful enough to sustain true sobriety.

She encounters a ritual wherein the storyteller's sole aim is to share the raw and open truth of their personal experiences. In its ideal form, such a story is told simply, without ego, embellishment, or comparison to others. I'm calling this ideal *lighthouse storytelling*, and I've become convinced that the surprising power of this practice makes it one of our best tools for building the cultures of the future.

Jamison's experiences with lighthouse storytelling in AA teach her the limitations of the Western humanism she was raised with, which casts meaning-making as a personal project, and shows her the beauty of relating communally, which allows meaning to be mutually created and self-referential (*"this is meaningful to me because it's meaningful to you"*).

As she sets out initially to write a book about how AA helped her get sober, Jamison finds herself treating her own life story as both subject and object, content and metaphor, deeply personal and widely shared. To her own surprise, it is not her recovery but rather *The Recovering* that becomes her redemption.

## The Recovering - The Fall

Like many people who gravitate towards writing, Jamison's natural disposition is compulsively self-reflective. It also tends to be strongly negative. She spends most of her teens and twenties in a cacophony of self-criticism, so overwhelmed by what she perceives to be her own flaws and errors that she often totally fails to hear the experiences and needs that others try to share with her.

"I was so self-absorbed there should have been a different word for what I was. Of course I would have loved that, if there had been a different word for what I was."

Over the years, she seeks to push back against the relentless assaults of this personal hell with an assortment of coping strategies including anorexia, cutting, and, eventually, a deep and destructive alcoholism.

Jamison deliberately lingers over these experiences in all their ugliness and shame, applying her talent for evocative scene-setting alongside an unflinching commitment to forcing us, her readers, to face reality alongside her. Reading the autobiographical sections of the book is immersive, moving, and painful.

"It was no wonder we got drunk so much; we just wanted a fucking break. Booze let me live inside moments without the endless chatter of my own self-conscious annotation. It was like finally going on vacation somewhere beautiful without having to pose for photographs the whole time."

"At a certain point we were on my bed and I didn't want to fuck him—but I was too drunk and too tired to figure out how not to fuck him, so I just lay there, still and quiet, while he finished. The situation would sharpen into awareness, in fleeting moments, and I'd think, This isn't what I want, and then it would dissolve into soft focus again."

"I wasn't using the word 'alcoholic' with other people, wasn't describing myself or my drinking that way, but those were the years when I started writing it in my diary, secretly, often during blackouts, syntax out the window: Is this what an alcoholic? It was as if a child just learning how to write had crawled inside my diary and called me by my name."

Jamison paints a rich and vivid portrait of American literary culture as it's embodied in the elite academic institutions she attends (Harvard, the Iowa Writer's Workshop, and

Yale). She gives a nuanced account of the ways in which that tradition has long valorized and enabled its heroes' self-destructive behaviors in the name of creative inspiration.

The writerly subculture surrounding her fits hand-in-glove within our humanist fixation on individuality, and both of them matched her natural self-obsession perfectly.

Her vision of creative success was intimately intertwined with her hunger to be, or perhaps to become, a special person. More specifically, the kind of success she always yearned for centered on producing something genuinely novel.

"In recovery, years later, when someone described self-loathing as the flip side of narcissism, I almost laughed out loud at the stark truth of what she'd said. This black-and-white thinking, this all-or-nothing, it was cut from the same cloth. Being just a man among men, or a woman among women, with nothing extraordinary about your flaws or your mistakes—that was the hardest thing to accept."

The book beautifully conveys the emotional texture and unresolvable tension of how it feels to believe in these cultural ideals of individuality, exceptionalism, and self-destructive genius. By her mid-20s, as her drinking has spiraled completely out of control and dominates every aspect of her life, it is clear to her that finding alternative stories to buy into is literally a matter of life and death.

As part of her all-consuming quest for a viable way of being in her world, Jamison does her entire PhD on American writers who struggled with, wrote about, and overcame (or didn't) their demons of addiction.

"I was moving between the worlds of graduate school and recovery, straddling the powerful rifts between their competing imperatives: Think harder. Don't overthink it. Say something new. You can't say anything new. Interrogate simplicity. Keep it simple. Be loved because you're smart. Be loved because you are. My dissertation was reckoning with a question I hoped might bridge these worlds, examining authors who'd tried to get sober and exploring how recovery had become part of their creative lives. It wasn't criticism as autobiography, exactly, so much as speculative autobiography—trying to find a map for what my own sober creativity might look like."

In this passage, Jamison reveals that she has already internalized the fundamental logic of lighthouse storytelling. To chart her own path to sober creativity, she doesn't need to invent a brilliant new solution to her intractable problems. Rather, she needs to hear personal stories from people who have gone through those same crucibles, become

changed, and emerged.

## The Recovering - Whence Redemption?

One easy criticism to level at the book is about the many privileges Jamison enjoys in her life. She has an essentially white and middle-class experience of addiction. Her story would have looked very different if she had been born brown or poor (among other things).

Jamison is well aware of this. She starts out skeptical of her own fit within AA, with its old clichés and unhip pastiche of struggling humanity, just as any intellectual young American of privilege would expect to be. But, and herein lies the heart of the matter:

"At that meeting, I was painfully aware of how much I had, and how much I hadn't lost. I was wary of how anything I shared might come across to others in the room, people who were struggling with so much more: the woman fighting for custody of her kids; the guy who'd been in and out of shelters for almost a year, but had finally gotten a job at the pizza parlor in town. How could comparing my addiction to theirs seem like anything but a misunderstanding of what they'd suffered? I didn't want to suggest I've been through that too, with my very presence—when of course I hadn't. My story was contoured by desire more than loss.

But I was surprised by the ways other people sought commonality, and at a certain point I realized I was the one projecting difference by assuming others felt it. Believing in what we shared didn't have to make me blind to what we didn't. Resonance wasn't the same as conflation. It didn't mean pretending we'd all lived the same thing. It just meant listening. [...] We weren't there to assume or insist on perfect correspondence; we were there to open ourselves to the possibility of company."

Analogously, many readers may struggle to find sympathy for the sad white girl drinking herself to death because she's not being recognized as a creative genius. And that's a completely valid reaction.

But if you are somebody who needs to hear what the book has to tell you, then the fact that the particulars of her life may not align with yours will not matter.

The fundamental patterns and demands of her addiction and recovery, which she

chronicles exhaustively, mercilessly, and in some measure also lovingly, will speak powerfully to you.

In the sharing of her journey towards recovery (her recovering, never to become a past-tense success until the day she dies), you will find communion.

Over the course of working on the project, Jamison realizes that her own goals have evolved. More than being simply about her own sobriety, her true goal with the book becomes to create a "Twelfth Step work," which alludes to AA's (final) Step Twelve: Having had a spiritual awakening as the result of these steps, we tried to carry this message to alcoholics, and to practice these principles in all our affairs.

"Googling the phrase 'just another addiction memoir' yields several pages of results, mostly blurbs insisting that a certain book isn't 'just another addiction memoir.' [...] This insistent chorus reflects a broader disdain for the already-told story, and a cynical take on interchangeability: the idea that if we've heard this story before, we won't want to hear it again. But the accusation of sameness, just another addiction memoir, gets turned on its head by recovery—where a story's sameness is precisely why it should be told. Your story is only useful because others have lived it and will live it again."

For it is always the case that nobody else has ever lived a life quite like ours; and also, that all of our triumphs and all of our tragedies have played themselves out uncountably many times before, as part of the grand shared human experience.

"'We tell ourselves stories in order to live,' Joan Didion wrote, and at first I took her words as gospel: Stories help us survive! But eventually I realized they were more like an admonition—a suggestion that there was something compromised and shameful about our dependence on their false coherence. When Didion wrote, 'I began to doubt the premises of all the stories I had ever told myself,' I understood her skepticism as an accusation: trusting stories was naive, a refusal to confront actuality in all its senselessness.

But in recovery, I started to believe again that stories could do all the things Didion had taught me to distrust, that they could lend meaningful arcs of cohesion; that they could save us from our lives by letting us construct ourselves. [...]

Recovery reminded me that storytelling was ultimately about community, not self-deception. Recovery didn't say: We tell ourselves stories in order to live. It said: We tell others our stories in order to help them live, too."

## The Current Landscape

I am not the first person to say any of this. Our culture is already moving in these directions.

### GENDER

[This NYT essay by Elena Ferrante](#) captures these themes in today's gender politics beautifully. She writes, *"There is one form of power that has fascinated me ever since I was a girl, even though it has been widely colonized by men: the power of storytelling,"* and closes with a call to action: *"The female story, told with increasing skill, increasingly widespread and unapologetic, is what must now assume power."*

On a personal level, I have watched three (women) friends over the past five years who have set out to write about certain vexing and rich elements of modern life - sex, feminism, the first-generation experience - and who have found, to their own chagrin and surprise, that the stories they're trying to tell keep demanding that they write themselves in.

In each case, their writing was originally intended to be third-person and/or fictional. And in each case, it became clear that an essential piece of the story's true potential lay in foregrounding the author's own personal journey, both with the subject matter and in the writing.

### IDENTITY

Over the past decade, the progressive left has developed innovations in identity culture at an astounding rate. Broader understanding of intersectionality, along with evolutions like fourth-wave feminism and Black Lives Matter, are meaningfully improving the tools people have for articulating their own experiences and organizing in solidarity around them. Identity-based culture is in fact so powerful that it has had various unintended side effects, including toxic call-out culture ([1](#), [2](#), [3](#)) and the resulting [ascendancy of "anti-PC" sentiment](#) across the political spectrum.

Some of the best parts of these identity cultures draw from lighthouse storytelling. Because of the patriarchal and white supremacist structures that are the bones of our modern society, identity-based experiences often create some of the most difficult and



essential parts of our selves. Many of my friends who are women and/or of color have found deep healing through sharing stories within identity-based affinity groups. Nothing can quite replace the feeling of speaking your truth to a group of people who really get it.

And some of the worst parts of these identity cultures are based on taking these principles too far. Our externally visible, socially constructed identities are powerful forces shaping our lives, but we are all complex, multi-faceted creatures. The potency of these new tools we have to define affinity based on identity can blind us to other, older ways of seeing commonalities and articulating shared life journeys.

## Don't Invent the New, Remember the Old

Many of us, like Jamison, have internalized the Western idea that progress comes from novelty, and that powerful stories for the future must therefore be *sui generis*, new in themselves.

If only we can tell the best story, the right story, the one that will pierce the veil of confusion and apathy that has descended upon our society and culture - then, we believe, we will finally ignite the passion that so many people find themselves missing so keenly.

I no longer believe in this theory of change. We must root ourselves back in the ancient forms of our human universals, the ones that are deep enough to provide common ground beneath the sound and fury created by our increasingly severe differences of circumstance and perspective.

Jamison's book succeeds because she is able to become an exemplar of a struggle and a triumph (addiction and recovery) that is both modern and ancient, and to use the power of her personal story to become a beacon, illuminating the path towards triumph for those who walk behind her.

And the fact that the book will succeed in its true aim only for those readers who find themselves enmeshed within that same struggle is in fact the very point.

## An Identifying Audience

One of the most powerful ways most humans learn is through hearing the stories of

others who are ahead of them on a shared path. One of the most powerful ways most humans teach is by sharing their own stories with people who are stuck a few steps behind them.

Any given personal story won't resonate with most people; that's an inextricable part of why they can be transformative for people attuned to their particular wavelength.

Support groups like AA are built around helping storytellers find the identifying audiences they need to discover the next step in their self stories. Thankfully, one of the best features of our Internet era is how much more quickly and broadly audiences and storytellers can find each other.

We cannot hope to reverse the cultural fragmentation that our technology is creating. We should at least seek to make use of it, to build structures that help people come together around sharing their individual stories of participating in the common struggles of being human in this world.

## Why Does This Matter So Much?

Lighthouse storytelling, applied in the right context, has a unique capacity to create and sustain strong communities.

To find meaning, we all need a story of self that somehow both centers us and subsumes us within a larger whole. We can build that kind of self-story collaboratively when we embed within communities of others who find themselves aligning around the same underlying human universal at the same moment in time.

Lighthouse storytelling is a democratic and populist approach. Its prescription for meaning-making is not genius invention, but rather mass participation.

In my reading and thinking about where we might turn to re-energize the community as a source of shared and co-created meaning, lighthouse storytelling is the best lead I've found.

And over time, in my writing and thinking, I hope to convince you, dear reader, that the power of this framework provides a crucial reason for optimism in the face of the overwhelmingly complex structural changes we need in order to build a just and sustainable future.

Any path forward will, after all, need to resonate deeply for people on both sides of what we perceive to be our current divides.

If we find the right common bedrock, we can build on it a world in which each person can see themselves, both in the glory of their individuality and in the comfort of communion.

May we all seek, find, and become bright beacons of old truths lived in fresh ways.

# The Red and the Black by Stendhal

Napoleon, in conversation which he once had with Goethe on the nature of Tragedy, expressed the opinion that its modern phase differed from the ancient, through our no longer recognizing a Destiny to which men are absolutely subject, and that Policy occupies the place of the ancient Fate [*La Politique est la fatalité*]. This therefore he thought must be used as the modern form of Destiny in Tragedy—the irresistible power of circumstances to which individuality must bend.

—Hegel<sup>1</sup>

Ancient political science understood regimes by the rule of the one, the few, or the many. A king reigned or oligarchs held sway unless the people got their say. Ideally, the community would mix these three together to form a middling element within society. In this golden mean, the few would mediate rulers and ruled through their wisely ordering the community towards the common good. However, this account arose in premodern societies that emphasized human inequalities. Perhaps, a modern account with new terms is needed to emphasize human equalities. Rather than kingship, oligarchy, or polity, modern political science understands regimes as aristocracy versus democracy. Politics is discerned not by rule of the one, few, or many, but by the rule of the greatest versus rule of the greatest number. Equality, not quality, is the ruling principle of modern society. This regime change of aristocracy to democracy means a change in hearts, one with levelled loves.

Innate human dignity was inconceivable to pagan antiquity. Thus, Alexis de Tocqueville writes, “it was necessary that Jesus Christ come to earth” to have “understood that all members of the human species are naturally alike and equal.” This religious transvaluation had latent effects. In Christendom, social structures altered over centuries until a democratic revolution commenced. Then, premodern aristocracies fell to democratic nations. “Christianity,” having “rendered all men equal before God,” enabled “all citizens equal before the law.”<sup>2</sup> This religious understanding turned into legal standing. Equality, no mere abstraction, is a social condition: a “sentiment du semblance” or “feeling of likeness.” While beliefs and structures of natural hierarchy dwindled, a crisis in mores began. Aristocratic souls intuit longings for the good and

1 Georg Wilhelm Friedrich Hegel, “Part III: The Roman World,” *The Philosophy of History*, preface by Charles Hegel, trans. J. Sibree (1899), introduction by C.J. Friedrich (Mineola, NY: Dover Publications, 1956), 278.

2 Alexis de Tocqueville, *Democracy in America* (1835, 1840), translated, edited, and introduction by Harvey C. Mansfield and Delba Winthrop (Chicago: The University of Chicago Press, 2002), 2.1.3, 413; “Introduction,” 11.

beautiful, yet democratic souls focus on material advantage. Human equality may seem self-evident, but noble and ennobling truths do not.

Less hierarchy means less moral nobility. Aristocrats mediated rulers and ruled. Democrats likewise need institutions mediating the state and equal individuals. De Tocqueville found in 1830s America this self-rule in business, religion, local government, and associations. This civil society educated isolated subjects into fellow citizens. But this ecology is fragile. Equality gave public opinion tyrannical sway democratic nations. “Founded on the ordinary order of things,” it brings masters and servants “to the common level” in sharing an “imaginary equality” despite “the real inequality of their conditions.”<sup>3</sup> But their bodily resemblance entails their spiritual alienation. Gone are old oaths binding loyal lords to their lands and men. These equals are connected on a purely financial basis. Further, an abstract love of universal humanity usurps embodied loves of particular people. And the individualism motivating pursuit of gain battles with voluntary neighborliness.

Restless democratic souls take great effort to fulfill their loves and rebuild social forms. As equals, we must exercise our liberty well. The nature of democracy is a state of nature, but its art renews the social contract. To operate so upon a body politic is *artfully* performed. Disenchanted by Bourbon France, de Tocqueville searched the United States for this art. Also shaken, Stendhal similarly worries public opinion tyrannizes democratic France. Its mention bookends *La Rouge et Le Noir* (1830). Chapter One, “Small Town,” features a Hobbesian epigram: “Put thousands together / Less Bad, / But the cage less gay” to describe the “prettiest” provincial town, Verrières. Many people put together is less worse than the alternative, yet the collective cage is less gay. In the new regime, fewer people are happy.<sup>4</sup> Monsieur de Rênal, entitled thanks to Napoleon, is mayor of a town housing a noisy nail factory: an image for the levelling of old things. What has it built?

Of French parochials, Stendhal notes, “these wise fellows wield an incredibly wearisome *despotism*.” Indeed, “this wretched word makes small towns unlivable” for those “successful in that great republic we call Paris.” French industrial countryside is unlivable to worldly Parisians. What despotism is wielded? “The tyranny of opinion—and such opinion!—is every bit as *idiotic* in the small towns of France as it is in the United States of America.”<sup>5</sup> Comparing old world to new, Stendhal diagnoses idiotic “tyranny of opinion” as a new despot. Its final mention goes: “The reign of public opinion” creates “inconvenience,” for while “it secures liberty, it meddles with what it has nothing to do with—private life.” And, “Hence the gloominess of America and England.” To “avoid infringing on private life,” Stendhal invented Verrières along with a bishop, jury, and

3 de Tocqueville, *Democracy in America*, *ibid*, 2.3.5, 550.

4 Stendhal, *The Red and the Black: A Chronicle of 1830*, translated by Burton Raffel, introduction by Diane Johnson, notes by James Madden (New York: Modern Library, 2003), 3.

5 Stendhal, *The Red and the Black*, *ibid*, Bk. 1, Ch. 1, 6.

court in Besancon, “where he has never been.”<sup>6</sup> Public opinion reigns in democratic towns all over.

Despots of the *ancien regime* were nobles and clerics lording it over everyone else, the old red and black, that is. The new despot, “wearisome” and worrisome, is public opinion. (And, “in the countryside,” the “husbands are in charge of opinion.”) Excessive, idiotic opinion is present in France and America. Its tyranny rules petit-bourgeoise, yet secures liberty. But this tool meddles with things properly outside it: private life, where passions reign and heartstrings play. “Genuine passion,” Stendhal says, “is egotistic.”<sup>7</sup> There, our selves are free to stride and strut, if private life remains sacrosanct from moeurs of political society. Yet, in a democracy, is it so simple? Private life should be distinct from liberty, which is political, and opinion, which is public. Opinion, from provinces to Paris, is a tyranny over private life. What can private life do when it cannot escape weary, despotic, wretched, unlivable, public opinion, a tyranny over freed people? It can revolt.

Note Stendhal mentions America both times. Gloomy England is substituted for France, a switch whose significance suggests the Republic of Paris is not so free of “such opinion,” nor mostly any place in Europe and America. Perhaps public opinion tyrannically took away private life’s liberty. Thus, passion is enfeebled. Feeling a “religious terror” at the “image of democracy” seen in America, de Tocqueville prophesies degraded souls without great longings. The perpetually returning passion for wealth made Americans monotonous. They are not alone. All men resemble each other more by pursuing the same object of gain.<sup>8</sup> Passion for advantage has overtaken love of honor. *Honneur* denotes the esteem, glory, and consideration a man obtains from his peers, and the rules for how he obtains these goods.<sup>9</sup> The ways to win esteem, glory, and consideration from peers belongs to aristocratic men. Honor means moral judgements not of all but only some men.

Specifically, it sanctions violence for such peers. The general, permanent human interest is to forbid murder, yet some men have particular, temporary interests to excuse or revere it. Honor is that “particular rule” of “a particular state” for how a people or class “distributes praise or blame.” In its preeminent world historical form, medieval aristocratic honor depended upon actor, typically, not actions. As war created feudal aristocracy, medieval nobles glorified valor first. Yet, Americans call “noble and estimable ambition” what medievals had named “service cupidity,” ascribing “blind and barbaric fury” to that “conquering ardor and warlike humor” of daily combats. Instead, democratic souls honor all honest work equally, thanks to “public opinion,” the natural sovereign interpreter of the honorable. As dissimilarities and inequalities of men created

6 This footnote does not appear in modern editions. Stendhal, footnote, *The Red and the Black: A Chronicle of 1830*, trans. Horace B. Samuel (New York: E.P. Dutton and Co., 1916). See [online](#).

7 Stendhal, *The Red and the Black*, trans. Burton Raffel, *ibid*, Part 1, Chapter 21, 126.

8 Alexis de Tocqueville, *ibid*, “Introduction,” 6, 13; 2.3.17, 588.

9 Mansfield and Winthrop (eds.), Footnote 1, de Tocqueville, *ibid*, 2.3.18, 589.

honor, it weakens as differences are effaced, to disappear with them.<sup>10</sup> But de Tocqueville has a prescription.

Americans have “so many ambitions” yet “so few great ambitions.” After a revolution, the aristocratic spirit subsists in confused desires because old mores wain. Equality allows all the shot to succeed at everything, but few achieve great things at anything due to competition. Aristocratic ambition is extensive in objects but fixed in its pursuit, whereas democratic ambition is narrow in objects but limitless in pursuit. Democratic souls love unlimited success more than definite glory. Thus, the greatest fear amid “the small incessant occupations of private life” is that “ambition will lose its spark” and “greatness,” with “human passions” both “appeased and debased,” that daily “the social body” is “more tranquil and less lofty.” To remedy this diagnosis, contemporaries should have a “vaster idea of themselves and their species.” Humility is unhealthy. Pride, vicious in aristocracy, is medicinal in democracy.<sup>11</sup> But, *as democracy disorients honor, does pride disorient love?*

---

Stendhal explores this possible trouble of prideful love as a potential passion in vain. Love of honor in a bourgeois society is rechanneled when society is pacified after revolutionary violence. De Tocqueville found industry and rights to have replaced honor, and likewise, Stendhal mentions opinion and liberty to have invaded private life. Honor is necessarily public. Its motive is prideful. But what happens to the passions of private life when swayed by passions of public life? Stendhal knew something about the *soul* which undergoes this democratic transvaluation. Referring to “this last great psychologist” who helps paint “the portrait of the free-spirited philosopher,”<sup>12</sup> Nietzsche quotes Stendhal in his own defense of *De L’Amour* (1822): “To be a good philosopher, one must be dry, clear, free of illusions. A banker who has made a fortune has something of the character needed to make philosophical discoveries, that is to say, to see clearly into that which exists.”<sup>13</sup>

Philosophy means seeing clearly into what is, and a successful banker can partly do that. This democratic personage peers beyond courts and newspapers “waxing eloquently about brilliant chimeras,” for, Stendhal laments, “the more public opinion will

10 de Tocqueville, *ibid*, 2.3.18, 590-591, 594, 597, 599.

11 de Tocqueville, *ibid*, 2.3.19, 599-604.

12 Such a type knows happiness and virtue cannot be used as arguments; that unhappy evil could be a part of the essential nature of human existence; that a man is counted as strongly spirited for how much truth he can tolerate; that the wicked are the ones who are happy; and the pain and cunning that makes a philosopher versus the ease and manner that makes a scholar; and to be a philosopher one must be dry, clear, and free of illusion.

Frederich Nietzsche, *Beyond Good and Evil: Prelude to a Philosophy of the Future* (1886), translated and edited by Marion Faber, with an introduction by Robert C. Holub (Oxford: Oxford University Press, 1998), Part II, Para. 39, 37-38; see original [online](#).

13 Stendhal, translation found in Faber, “Explanatory Notes,” Nietzsche, *ibid*, 184.

become the queen of France, the more there will be hypocrisy and cant,” which “is one of the inconveniences of liberty.”<sup>14</sup> Public opinion and political liberty have evil twins in hypocrisy and sophistry. The banker can partly see these things as what they are. This democratic man can take advantage of hypocritical cant. Thus, he successfully wins his fortune. Fortunate and unfortunate bankers are distinguished much like successful and unsuccessful lovers. To see market signals, win clients, invest and loan prudently, and keep coffers filled requires a man who understands hypocrisy and cant, even by trading in these children of public opinion. This bourgeois undertaking revels in democratic debasement.

That is a Don Giovanni of finance. But what of the Don of romance? Genuine passion is egotistic, not altruistic, for Stendhal. Thus, love as a passion is not about the beloved, but the lover, and not his mistress, but his pleasing idea of her. Hellenism and Christianity, in contrast, accounted for love as an activity. Ancient Greeks knew four kinds – storge, eros, philia, and agape – loves filial, romantic, friendly, neighborly, and divine.<sup>15</sup> Christians elevated love into charity: the love of God given to men called to love of neighbor. All loves involved affections of the heart, but also benevolence in the will, along with unions of relations. The truth, “Greater love hath no man than this, that a man lay down his life for his friends,” concerns not just inward passions, but real merciful actions by grace ennobling “servants” into “friends,” whom “charity edifieth” even whilst we still “see through a glass, darkly.”<sup>16</sup> But an emotivist account of love reverses its intentionality.

The sophisticated banker and the clarifying philosopher see past chimeras, whereas the successful lover creates his own ideals. Ideal love is a passion, and passion is biological.. The love for kith and kin, for example, has particular objects, but is

14 “5 ° Contre l'opinion des femmes : la philosophie allemande cherche toujours à émouvoir le cœur et à éblouir l'imagination par des images d'une beauté céleste. Pour être bon philosophe, il faut être sec, clair, sans illusion. Un banquier qui a fait fortune a une partie du caractère requis pour faire des découvertes en philosophie, c'est-à-dire pour *voir clair dans ce qui est*;

“Ce qui est un peu différent de parler éloquemment de brillantes chimères.

“Plût à Dieu que tous les hommes fussent des anges ! Alors, plus de juges prévaricateurs, plus d'hypocrites, etc., etc. Voyez les journaux: ils vous disent que nous sommes loin de ces chi mères. Plus *l'opinion publique* deviendra la reine de la France, plus il y aura d'*hypocrisie* et de *cant* ; c'est là un des inconvénients de la liberté.”

Stendhal, Letter to Sutton Sharpe, 24 October 1829, “Lettres A Ses Amis,” *Correspondence Sédite*, introduction by Prosper Mérimée, Vol. 2 (Paris: Lévy, 1855), 87, [Google Books](#). Free translation.

15 “It is nice to have arrived at all this by something which began in Agape, proceeded to Philia, then became Pity, and only after that, Eros. As if the highest of these, Agape, had successfully undergone the sweet humiliation of an incarnation.”

C.S. Lewis, “To Dom Bede Griffiths, O.S.B.: *from the Kilns*,” 24 September 1957, *Letters of C.S. Lewis*, ed. Warren Hamilton Lewis (London: Harcourt Brace, 2003): 469-470, 470.

16 King James Bible, John 15:12-17, 1 Corinthians 8:1, 13:1-13.



universally held: “earnest speeches of foreign patriots” have “sentiments which inspire exactly similar statements elsewhere.” But its manifestation differs due to circumstances. “In Constantinople” and “all barbaric countries, this blind and exclusive partiality for one’s own land is a fury which demands blood,” whereas, “among cultured peoples it is a pained, unhappy, anxious vanity, that turns at bay on the very slightest provocation.” Patriotism is everywhere felt, but everywhere felt variously. Likewise, erotic love is universal, various in kind, yet so understandable as to be given “an exact and scientific description of a brand of madness very rare in France.”<sup>17</sup> Love is a rare madness vainly to be found in France.

All love leads to Rome, ideally, in its crystallization. This process conveys “the whole congeries of strange follies about the beloved” that get “regarded as true and beyond question.” The four kinds of love – passionate, mannered, physical, and vanity – can admit even “eight or ten distinctions,” yet while there exist “as many different ways of feeling” as “of seeing,” variations in terms alter not that “every variety of love” is “born, lives, dies, or attains immortality” according to “the same laws.”<sup>18</sup> These laws are metaphorically a journey. In travel from the indifference in Bologna, across the Apennines, then on road to perfect love in Rome, the journeying lover travails from admiration, to acknowledgement, then hope, and lastly, delight. But not all loves are equal. Stendhal clearly finds passionate love not only supreme, but to almost mystically justify living. But is this living not union between lover and beloved, but the man experiencing his desires satisfied?

Beauty is “a new potentiality for pleasure” from the mistress. Her beauty is the sum total fulfillment of all desires formulated about her. It transcends physical attraction as “the promise of a quality useful to my soul.” This fulfilled sum is the crystallization of the mistress in the lover. Here, the beloved is not an object of love, but its product. It encourages the lover to imagine hopeful desires with consequent satisfactions. These associations together constitute being in love, while boredom means the withdrawal of love. One must continue the game. The better lover, like Young Werther, preys upon panics and hopes to sustain crystallization. Fear of losing a pleasure ensures its desirability for vanity, but it must be by effort, otherwise love delights less. Needed is the swaying between opposites: hope and fear, conquest and loss, and war between lovers. Other desires adapt to “cold reality,” but in love desires rearrange to conform not to reality, but desire.<sup>19</sup>

Morals, politics, and religion intrude upon love to end it. The love which lasts continually entertains its mental fixtures, but the love which dies gains full satisfaction with its beloved. Marry the mistress and fall out of love: hence the distinction between Don Juan and Young Werther. Don Juan thinks love “is a feeling,” like “a taste for hunting,” whose

17 Stendhal, “Preface to the First Edition,” *Love*, translated by Gilbert and Suzanne Sake, with an introduction by Jean Stewart and B.C.G. Knight (New York: Penguin Books, 1975), 23, 25.

18 Stendhal, “Final Preface,” *Love*, *ibid*, 35; “Chapter 1: On Love,” 44-45.

19 Stendhal, “Chapter 11,” Footnote I, *ibid*, 59; Chapter 12, 60.

“craving for an activity” needs “incessant diversity of stimuli to challenge skill.” New prey is needed. Werther, however, always finds new ways to see his mistress in a thousand fold “magical visions.”<sup>20</sup> Ever promiscuous, Don Juan must possess new realities in new loves, but Werther stays true by ceaseless imaginings. Bankers and philosophers see past these realities, but the worthy lover entertains chimeras. They, like Don Juan, understand the external too well, but Werther ever strives internally in discovering fresh perfections in his beloved at every new turn. This is the passionate love AWOL in France.

Inconstancy is cursed with boredom. Passionate love is cursed with death and despair. Each type has its virtue, yet boredom entails death, a curse to both Juan and Werther. “More brains are blown out for love than from boredom.’ I can well believe it, for boredom strips away everything, even the courage to kill oneself,” Stendhal writes.<sup>21</sup> In the democratic age, boredom is stymied with ever new stimulations, however homogenous and bourgeois. But why do Frenchmen not endow their women with perfections they do not possess? And what occurs when they try? For the old religion, to make an ideal of the mistress is to craft an idol. In the new politics, the old idols were smashed. Revolution meant iconoclasm. So one must restlessly seek out new pleasures and enterprises. The banker attends the bourgeoisie, the philosopher their psychology, the lover their ennui. But what of the conqueror? What does Napoleon love?

Does reddening love relate to blackening war? Love “springs from Nature” ordaining “we feel pleasure” while sending “the blood to our heads.” Its growth “evolves” from “feeling” degrees of pleasure relate to the mistress’ perfections and “from the idea that ‘She is mine.’”<sup>22</sup> This crystallization can describe not just love of women, but love of glory, as well. Things were not always thus, that only Greeks sought independence, or just barbarians were furiously partial to their bloodied land, whereas cultured democracies had their passions honor public opinion. Once France had her revolution, terror, and empire. “My love for Napoleon is the only passion remaining to me,” Stendhal wrote in 1836.<sup>23</sup> He chronicles the European misfortune that Napoleon was miseducated in the old regime. Not having read enlightenment writers, he did not learn of “the strength which public opinion confers upon a Government,” for “the mixture of Catholicism and aristocracy” had “ironed out our souls for the past two hundred years.”<sup>24</sup>

Even George Washington would have hesitated to give Frenchmen too much liberty! Experience was nothing to a nation still fundamentally cherishing inherited feudal

20 Stendhal, “Chapter 59: Werther and Don Juan,” *ibid*, 209.

21 Stendhal, *ibid*, 211-212.

22 Stendhal, “Chapter 2: Concerning The Birth of Love,” *ibid*, 46.

23 Stendhal: “My love for Napoleon is the only passion remaining to me; yet it does not prevent my seeing his faults and the petty weakness with which he can be reproached.” Quoted in Roland Grant, “Introduction,” Stendhal, *A Life of Napoleon* (London: The Rodale Press, 1956): 1-4, 4.

24 Stendhal, *Life of Napoleon*, *ibid*, 9; *ibid*, 22-23.

prejudices. Yet Napoleon was no Solon. “Exceedingly ignorant of the art of government,” Bonaparte, bred on military ideas, judged France too weak for visionary ones. Confident in his superiority and hateful of political discussion, Bonaparte embodied Roman ideals, believing the worst evil was to be conquered, not badly governed: “he was too superb a general to be any good as a politician and legislator.” Afraid of the masses, he never planned, yet his political mistakes bore fruit. “The institutions” he founded “were always liberal,” easily perfected to “bring forth liberty,” and advantageous for France to “forget everything old,” since “Frenchmen needed to be cured of their respect for outworn ideas.” Stendhal gives this judgement upon his countrymen:

‘The French, Napoleon remarked, ‘are indifferent to liberty. They neither understand it nor like it. Vanity is their ruling passion. And political equality, which enables them to feel any position is open to them, is the only political right they care about.’ *Never has anything more true been said about the French people.*<sup>25</sup>  
[Emphasis added.]

People inhabiting mediating institutions can withstand the bipolar oscillation of public opinion. Rooted in religious mores, their restless souls can innovate in ways political, economic, and technological, all thanks to their institutional communion buffering them against this soft despotism. But if unrooted and unformed, their desire for social mobility opens their hearts to malformation. Public opinion is that potent. Such Tocquevillean points complement the remarks of Napoleon. Stendhal affirms his opinion: the French, indifferent to liberty, neither understand nor like it. Rather, their ruling passion is vanity, while political equality, the only right they esteem, opens their imaginations that any position is achievable. Vanity grounds love of equality. This love of mobility rules French hearts. Here, Nietzschean psychologist Stendhal shows the consequence of this transvaluation from the democratic revolution. He shows the problem of Napoleonic love.

---

In *The Red and the Black*, Stendhal shows how pride affects love within a democracy. In a bourgeois society, public opinion creates perverse incentives for prideful desire. Worried Monsieur Valenod, the other richest citizen of Verrières, will engage Julien, M. de Rênal redoubles his desire with Monsieur Sorel to engage Julien instead. Why did “this petty man” with “his petty fears,” Stendhal asks, “bring a man of spirit into his house” and not “someone with a servant’s heart,” whom he needed? Why does M. de Rênal not “know how to pick people?” This chapter, “A Civil Servant’s Sorrows,” begins with the Casti epigram that the pleasure to hold the head high yearlong is surpassed by necessarily enduring “certain fifteen-minute intervals,” that is, when the head must be bowed. La Revolution dreamed every man could hold his head high, but sometimes he must bow. Public opinion creates within the soul a desire for competition when there may not be any.

<sup>25</sup> Stendhal, *ibid*, 36-37, 40-41, 42.

But democratic souls have replaced aristocratic souls. Hiring a spirited man, Julien Sorel, this petty Verrières Mayor cannot properly run his household as custodian to public opinion. Usual nineteenth-century procedure “is that, meeting a man of spirit, a powerful nobleman, promptly” kills, exiles, imprisons, or humiliates him to death, yet, “here, by accident,” the man of spirit suffers not so. Greatly misfortunate, small French towns and “places governed by elections (like New York)” cannot escape from the Rênals. In middling cities, “these men” like de Rênal “shape public opinion,” for “public opinion is a frightful thing in a country thing that has a constitution.” Thus, in a democracy, “A noble-souled, generous man” who should be “your friend” but lives 500 miles away, “judges you by the public opinion of your city” as “formulated by fools” whom chance gave good birth, riches, and security. The lesson is “woe to the man who stands out from the rest.”<sup>26</sup>

Thrasyboulus’ lesson, “to murder the prominent men of the city,” is a democratic one.<sup>27</sup> Public opinion means, Madame de Rênal advises Julien, “Be sweet, be polite, don’t be nasty even to the most vulgar people,” for “they’re going to decide our fate.” It means, M. de Rênal tells his wife, “the strange condition of public opinion in Verrières,” which is to be “misled by the envious.” It entails, for fifty Besançon seminarians, “whom public opinion had made mindful of living reality,” the “Jacobinism” ready, “lying in ambush behind every bush.” (It is, for Stendhal, pressure from publishers to satisfy readers by having his characters talk politics, otherwise “they’ll cease to be the Frenchmen of 1830, and your book will no longer be a mirror.”) It is, per Monsieur de La Môle, “who and what must be crushed,” specifically “journalists, voters, public opinion—in short, youth and all who admire it” who “stupefy themselves with the noise of their empty words.”<sup>28</sup>

Already, Stendhal charges thrice that the mechanism of public opinion damages private life. Further, public opinion is accounted as things ranging from the courtesy due to the vulgar who decide fates, envious folk who mislead mayors, anticlerical Jacobinism, public pressure on authors, and journalists, voters, and youth stupefied by their empty noise. It is altogether a moral degradation that muddles men of spirit and elevates petty men. When M. Valenod befriends the *Congrégation* to become the ultraroyalist candidate, M. de Rênal decides to become a defected liberal. Competition creates an opposite mimicry, where principle easily shifts and honor can be bought. Fulcoz and Saint-Giraud converse about the causes for the troubles of provincial life. For Saint-Giraud, that “ghastly life” is a crime thanks to the deposed emperor. In decades hence, politicians will be smarter but not more moral while reformers and conservatives will remain seeking office.

26 Stendhal, *The Red and the Black*, trans. Burton Raffel, *ibid*, Part 1, Chapter 23, 140.

27 Herodotus, *The Histories*, The Landmark Herodotus, translated by Andrea L. Purvis, edited by Robert B. Strassler, introduction by Rosalind Thomas (New York: Anchor Books. 2007), 5.92ζ-η, 408-409.

28 Stendhal, *ibid*, Part 1, Chapter 20, 114; Part 1, Chapter 23, 149; Part 1, Chapter 28, 184; Part 2, Chapter 22, 361; Part 2, Chapter 22, 364.

Regarding “the Ship of State, everyone wants to steer, because the job pays so well,” he remarks, “But will there ever be some little spot that’s open to an ordinary passenger?” The open competition means the elites still grab power that remains closed to ordinary citizens. Saint-Giraud, “tired of the perpetual comedy” all “play, in Paris,” a play called “nineteenth-century civilization,” got “thirsty for goodwill and simplicity.” Yet the priests and aristocrats caused this apolitical man to become a liberal due to the lawsuits. Losing money and selling his house, Saint-Giraud wants “out of that inferno of hypocrisy and petty troublemaking,” and instead to seek countryside “peace and solitude” where only in France such things are “found, a fourth-floor apartment overlooking the Champs-Élysée.” *Only in central Paris can pastoral peace and quiet be had.* But, Fulcoz replies with angry sorrow, “All that wouldn’t have happened, under Bonaparte.” Yet, Saint-Giraud dissents.

“Everything I endure today stems from what he did.” This decent harmless man at 40 with 500,000 Francs cannot peacefully live in the provinces due to despotic priests and aristocrats. The Emperor is to be blamed, yet Fulcoz notes: “France never stood so tall, in the eyes of the world, as it did in the thirteen years of his reign. Back then, there was something grand about everything we did.” Napoleon won honor for France, yet he also sold it. As Napoleon was friendly to clerics and lords, Saint-Giraud says, he had to leave his property. “It’s Bonaparte” who had “put the Rênals and Chélons in power, and it’s them who gave us the Valenods and the Maslons.”<sup>29</sup> Napoleon granted aristocratic titles based on income, not birth or service. Honor could be bought, but the whole point of honor is to be higher than money. What is not for sale? “Nothing can so distinguish a man as a death sentence,” Mlle. Mathilde notes. “It’s the only thing one can’t buy.”<sup>30</sup>

Their conversation awoke Julien from his usual mental chimeras. After hearing them speak, Julien “silently swore never to abandon his beloved’s children; he’d leave everything in order to protect them, should priestly arrogance ever bring on a republic and persecution of the aristocracy.” Here his intellectual confusion is on full display, that priests would support a republic to persecute aristocrats, their bedfellow victims in the Reign of Terror. Yet further, that his passion is for his beloved, Mme. de Rênal, and her

29 Stendhal, *ibid*, Part 2, Chapter 1, 219-223. Mr. Jeffrey Smith lectures,

“The irony of the situation was that this new class of the newly rich owed their wealth to the man the Bourbons deposed, and specifically to the economic policies Napoleon had pursued after 1800. Requiring greater and greater tax receipts to fund imperial expansion, Napoleon had both ramped up domestic production throughout France and erected protectionist barriers throughout Europe against British goods. The state became wealthier because the economic actors throughout France became wealthier, including the workers, producers, and manufacturers in rural France, who were subsidized by a state that also helped to buy up their goods all the while facing weakened competition from abroad.”

Jeffrey Smith, “Stendhal’s Prophecy for Liberal Democracy: Thoughts on *The Red and the Black*,” lecture delivered at Saint John’s College, Annapolis MD, 17 February 2012.

30 Stendhal, *ibid*, Part 2, Chapter 8, 274.

children, is also confused. As Stendhal comments, “Real passion always thinks of nothing but itself. That is why, it seems to me, passions are so ridiculous in Paris, where your neighbor acts as if you’re always deeply attentive to him.”<sup>31</sup> Passion is always egotistic. It is essentially not selfless. But in his crystallization, Julien holds to his fantasies. This conversation reveals Napoleon, who sought aristocratic imperial glory yet whose institutions were perfected to liberty, preconditioned the content of public opinion reigning in Bourbon France.

The bildungsroman of Julien Sorel is straightforward. His journey begins among provincial nouveau aristocrats, then later among established Parisian nobility. In Verrières, his first beloved is Mme. de Rênal, and his first employer yet romantic roadblock is M. de Rênal. In Paris, his second beloved is Mlle. Mathilde de La Môle, and his second employer yet romantic roadblock is the Marquis de La Môle. After his first amour, he attends seminary in Besançon, and after his second amour, he attends his trial. This *son of a carpenter* knows the Latin New Testament without sanctity or understanding: “instead of these wise reflections, Julien’s soul, wandered in imaginary places. He’ll never make a good priest, nor a great administrator. Souls stirred like this are good, at best, for making artists. Here, Julien’s vanity bursts into the full light of day.”<sup>32</sup> His vanity is his art. His hypocrisy is his brush. And his personal life is his canvas. Will his paintings see the light of day?

Not fit for priesthood nor administration, he is best fit for artistry with his vain wanderings. Yet Julien is supposed to be a man of spirit, despite his “extraordinary ignorance” and “coarseness of his manners,” later receiving Father Pirard’s lessons in the diction of the Parisian nobility. His soul is a confused mixture of pride acquiring love but negating its flavor tasted. Because, “amid those sweetest moments” with Mme. de Rênal, he “remained victim of his bizarre pride” in “pretending to be a man accustomed to subjugating woman,” Julien “made unbelievable attempts to spoil what was loveable.” Unmindful yet remorseful, Julien had “the notion of *duty*” gripping “his attention,” afraid of remorse and scorn should he abandon his “ideal model” of himself. Stendhal concludes: “exactly what made Julien a superior being” prevented his “enjoying” his newfound “happiness.” A superior being of spirit, Julien is yet ignorant, coarse, but learning.

Victim to bizarre pride with illusions of virility, Julien keeps spoiling the loveable. Still, he must attend to duty, but duty to what? Stendhal says Julien is “like a young girl of sixteen, with a magnificent complexion, who before she goes to a ball” foolishly will “daub herself with rouge.”<sup>33</sup> Julien is a superior being whose makeup, his ideal model, inhibits his joining the happy few. Why? Well, his ideal life is modelled after Napoleon. Getting inspiration from the ancien regime, Mme. de Rênal looks to Catholicism, the black; Mlle. Mathilde looks to warrior ancestor, Joseph Boniface de La Môle, the red.

31 Stendhal, *ibid*, Part 2, Chapter 1, 224.

32 Stendhal, *ibid*, Part 1, Chapter 28, 184.

33 Stendhal, *ibid*, Part 1, Chapter 11, 36; Part 2, Chapter 1, 224-228; Part 1, Chapter 15, 82.

Julien looks to enlightenment ideas contra religion and military glory contra royalism. When M. Sorel finds Julien reading idly, he beats his son and throws *Le Mémorial de Sainte-Hélène* into a river. "There were tears in his eyes, less from the physical pain than for the loss of his book, which he dearly loved." Julien is in love, not with Napoleon, but his idea of Napoleon.

But Napoleon is not alone. "His repugnance" at proletariat life Julien got from "Rousseau's *Confessions*," the "only book that in any way helped him imagine the great world," which alongside a "Grand Army bulletin" collection "and *Memories of Napoleon on Saint-Helena*, completed his personal Koran." Ready to die "for these three books" and "never believe in any others," Julien deemed "every other book in the world a pack of lies, written by swindlers trying to get ahead." Despite his memorizing the Latin Vulgate and reading *On the Pope* by Joseph de Maistre, with "no more belief in the one than in the other," in his intellectual confusion, "together with his fiery soul, Julien possessed one of those stunning memories so often linked to stupidity."<sup>34</sup> His stupidity is the rouge he uses, the social convention that puts chains on man born free, for Julien in his solitary walks is a natural man with romantic dreams whose democratic covering debilitates his superiority.

Perhaps. It might be at least said, Julien Sorel feels, therefore Julien Sorel is.<sup>35</sup> Napoleon haunts not just Sorel, but the novel, which mentions his name over sixty times. Neither the red nor the black of aristocracy, the Catholic New Testament nor the reactionism of de Maistre, sway Julien, who is charmed instead by the red and the black of democracy, Bonaparte and Rousseau. How do these types fuse together inside his superior stupidity? Napoleon means the glories of imperial victory, while Rousseau means a sentimental education by experience. Thus the external heroic conquest turns inward into a spiritual struggle. Passion is rare in France, but in Julien? His nobility is not landed gentry, but spiritual. The aristocrat is the ultimate passionate man, but the Bourbon aristocracy, old and new, is devoted to vanity. All have vanity, but not all have passion. Whom do we imitate in a democracy, the New Testament and the Catholic Church, Joseph Boniface de La Môle and Joseph de Maistre, or Napoleon Bonaparte and Jean-Jacques Rousseau?

Democratic souls when in competition get envious at each as each seeks to be godlike. The nobility mistakenly try to retain its privileges by imitating the bourgeoisie, thus the aristocracy becomes democratic through its hatred of democracy. This ruling class covers its cheeks with rouge before going to the ball. There is hypocrisy for vanity, such as how Julien and Mlle. Mathilde spar, and there is hypocrisy for profit, the successful banker. Put the two together, the doomed selfish striving for passion and the

34 Stendhal, *The Red and the Black*, ibid, Part 1, Chapter, 4, 16; Part 1, Chapter 5, 19-20.

35 "In vain do you argue this point with me; I feel it, and it is this feeling which speaks to me more forcibly than the reason which disputes it."

Jean-Jacques Rousseau, "The Creed of a Savoyard Priest," *Émile, or Education* (1762), translated by Barbara Foxley, M.A., (New York: E.P. Dutton, 1921), [online](#).

public climb to recognized success, and what one gets is Julien. Hypocrisy is how he makes his Napoleonic conquests of society. His duty to hypocrisy is his passion in vain amidst the tyranny of public opinion after Napoleon. Further, pride overtakes and corrupts romantic love without traditional hierarchies meant to channel pride. Our psychologist shows the tragedy of Julien Sorel is born of the perverse use of pride enabled by democracy.

---

Julien Sorel holds to new ideals of excellence, yet these ideals are historically acquired. New idols replaced old ones even in the countryside where Napoleon aided the two estates. Since his fall, “provincial manners rigidly suppress anything that smacked of gallantry.” As people fear dismissal from office, rogues seek support from the ultraroyalist Congregation of the Holy Virgin. Yet, for all types of persons, “hypocrisy has made excellent progress,” even among the liberals. As “boredom grows and grows,” only reading and farming remain available pleasures.<sup>36</sup> In this setting, what are upstarts to do? One way is to be class-conscious, which Julien had already gleaned from Rousseau. His knowledge of his societal status is tied up with his hypocrisy. “At church, all I ever see is God,” Julien tells his father, “a bit hypocritically, yet a fair and decent maneuver,” to avoid a paternal beating. Religion is a hypocrisy for justified social maneuvers, and Julien is a pro at it.

“And what would I be paid?” Julien, financially ambitious, asks his father, for he is aware of his petite-bourgeois standing: “I’m not interested in being a servant.” But, M. Sorel asks, “Why would I want my son to be a servant?” Clearly, his aggravated father wants his son to advance. Yet talk of mobility may go only so far. “But who will I be eating with?” Julien asks, embarrassing his father fearful “he might say something rash.” His father knows some topics are not to be broached, especially class politics. Objecting to his expected domestic status, for one of Napoleonic ambition and Rousseauian sensibility, he lacks honest self-awareness to strategically advance. Affectionless for the children adoring their good tutor, Julien banished household boredom but “his mind was elsewhere” in his “hatred and horror for the high society” where “he had been admitted—at the foot of the table,” true, which would “explain the hate and the horror.”<sup>37</sup> He detests as he climbs.

His prideful desires even inhibit his taking advantages of opportunities for his success. In refusing the marriage proposal by the Rênal chambermaid, Mademoiselle Elisa, whose “coming into some money” would provide the dowry for an enjoyable middle-class lifestyle. Julien refuses the offer, for “my secret ambition” is “to make my fortune.” But now he could have his fortune, yet this felicitous domestic offer is not the fortune he envisions. His duty for vanity means refusal to marry within his class. Money alone will not satisfy his desires. His refusal shows his pride conflicts with his luck. So what does it mean to be fortunate? How does a banker gain his fortune to become a

36 Stendhal, *The Red and the Black*, *ibid*, Part 1, Chapter 8, 42.

37 Stendhal, *ibid*, Part 1, Chapter 5, 18-19; Part 1, Chapter 7, 33.



savvy philosopher? But, Stendhal claims, “there is no need to prophesy terrible things for Julien.” His excuse, in slandering Mlle. Elisa, was “nicely created language” of “a crafty, prudent hypocrite,” which “at his age” was “hardly wickedness.” If the reader detects sarcasm, he should.

But Julien displayed integrity in his hypocrisy. Having sworn to say only what appeared fallacious, he was still amateurish in speech. Because “gestures” he used came from “country folk, he had never had the opportunity to watch the great models.” However, “given the chance to mingle with these gentlemen, he would become equally polished in both gesture and word.”<sup>38</sup> He would learn from the great models for how to be a better, more gentlemanly, hypocrite. Still, Julien can only learn so much by imitation. He remains impulsive for a fight: finding “all the apparatus of a provincial financier who considers himself a Don Juan” had “only made him think, once again, of the cudgeling he owned the man.” In this scene, Julien understands, “He had acquired the manners, but not the heart of his social status. For all his hypocrisy, so frequently practiced, he felt a heavy tear rolling down his cheek.”<sup>39</sup> And practicing hypocrisy is the paradigmatic virtue of Julien.

Finally leaving his father, Julien “thought that making a stop at church would be important to his hypocrisy.” At this first mention of his hypocrisy, Stendhal asks, “that word surprises you? Before arriving at so horrible a conclusion, the young peasant’s soul had traveled a long road.” He had seen dragoons from the Italian campaigns, heard tales of battle from the old surgeon-major, witness a church built, and heard of reactionaries punish liberals. When the surgeon-major died, Julien concealed his talk of Napoleon and declared his wish to become a priest, for, “Those who were on God’s side had won.” Stendhal exclaims, “Who could have imagined that this girlish face, sweet and pale, concealed an eradicable determination to risk a thousand deaths rather than fail to make his fortune!” And “making his fortune meant, first of all, leaving Verrières.” For his hometown “chilled his imagination.” His Napoleonic imaginings encourage his priestly vocation.

Since early childhood, Julien experienced “moments of exaltation.” Dreaming of delights to be found among Parisian ladies, “he would know how to attract their attention by some brilliant act.” These grandiose dreams were due to Napoleon. Why should not a Parisian woman love Julien like the Mme. de Beauharnais loved Bonaparte “while he was still poor.” Almost every hour for years, Julien recited: “Bonaparte, an obscure, poverty-stricken junior officer, had conquered the entire world with his sword.” This faith consoled him for his misfortune and doubled his pleasures. The church construction and the local justice of the peace inspire Julien, that since clerics wielded greater power than judicial officers, he chose black over red: “So the thing is: to become a priest.”<sup>40</sup> Or rather, his idea of the red gave him his first idea ever, to join his idea of the black. This

38 Stendhal, *ibid*, Part 1, Chapter 7, 43-45.

39 Stendhal, *ibid*, Part 1, Chapter 22, 131-132.

40 Stendhal, *ibid*, Part 1, Chapter 5, 22-23.

faith in Napoleon allows Julien to imagine greater possibilities for his life than working like his father did.

Napoleon was a godsend, so Julien tells Mme. de Rênal. "God obviously sent Napoleon for the youth of France!" However, his fall from power gives cause for worry. "Who can take his place?" he wonders out loud. Without Napoleon, what will other miserable young men do, who, though richer than Julien, who have enough money for good educations, but not enough, at 20, to bribe the right people for a career? Julien sees Napoleon not as an inspiration for conquest, but as an example to emulate for boys to rise from obscurity so to master the world by their deeds alone. In contrast to the aristocratic model of heroic inspiration, where ancient deeds are emulated to inculcate modern virtue,<sup>41</sup> now democratic passions are the new fields of battle. Can Julien win the war there? "Whatever we do," Julien deeply sighs, "this fateful, imperishable memory will forever keep us from being happy!" The unhappy "us" can refer to the lovers, or all Julien Sorels.

Probably, Julien fears he and his beloved will not be one of the happy few in the modern era. *The memory of Napoleon is the reason for why so few are happy*. No one can replace his glory. But souls can try to feel like he did, but in vain. The reaction of Mme. de Rênal explains why it is so. At this speech, she frowns with "a cold, disdainful air," since "that way of thinking" was "better suited to servants." However, thankfully, since she was "raised with the idea that she was extremely rich, it pleased her to assume that Julien had been, too." Her faux aristocratic pride is integral to her *idea* she was nobly born into riches, and she assigns this *idea* of inherited wealth to Julien. Of course, Julien was not born so. However, love covers multitudes of sins, or at least factual inconveniences. Thankfully, "she loved" Julien "a thousand times more than life itself, and never gave a thought to money."<sup>42</sup> Only the wealthy can give no thought to money. But all may mistake love over life.

This triangle of the lover, the beloved, and an ideal we see again and again. Consider our hero on a Rousseauian solitary walk, a way to escape decadent society

41 "In a democracy where men already think they are weak, they are too open to theories that teach that they are weak, which, by making individuals think that controlling action is impossible, have the effect of weakening them further. The antidote is again the classic, the heroic—Homer, Plutarch. ... Churchill was inspired by his ancestor Marlborough, and his confidence in his own action is inconceivable without the encouragement provided by that model. Marlborough said that Shakespeare was essential to his education. And Shakespeare learned a large part of what he knew about statesmanship from Plutarch. This is the intellectual genealogy of modern heroes. ... Tocqueville did not believe that the old writers were perfect, but he believed that they could best make us aware of our imperfections, which is what counts for us."

Allan Bloom, *The Closing of the American Mind: How Higher Education Has Failed Democracy and Impoverished the Souls of Today's Students*, forward by Saul Bellow (New York: Simon & Schuster Inc, 1987), 255-256, [online](#).

42 Stendhal, *ibid*, Part 1, Chapter 17, 88.

and return to our true nature. Julien standing on mountain rock gazes at the August sunlit sky. When the crickets finish singing, Julien is "surrounded by absolute silence." Looking over leagues of countryside, he watches sparrow hawks fly in quiet circles. Julien eyes these "bird of prey," as "the hawks' quiet, powerful movements impressed him; he envied their strength; he envied their utter isolation." Now comes his ideal. "That had been Napoleon's destiny. Would it be his, one day?"<sup>43</sup> Napoleon is in his mind the perfection of what the solitary walker sees: the utterly isolated but strong hawk. Julien wishes not to escape society, but rule it. That is his fortune. And he strategically triangulates his loves like a bird of prey, or so he imagines. This is a chimera. And the perfect lover always invents new ones.

Julien is hawkish at his social hypocrisy. His "whole congeries of follies about the beloved" which he regards "true and beyond question" is his own Napoleonic life, all in his head, which he connects to new prospects. Yet his pride, the *rouge*, is his ruin. First, Julien turned down the offer of marriage by Mlle. Elisa and her newfound wealth. Second, he turns down the offer by Monsieur Fouqué to begin together a lucrative timber business. He could make a great conquest in love by nuptials, or make a great conquest in riches by business. Yet he moderates not his longings. With the money, Julien would "be in a better position to become a soldier, or a priest, whichever would be better, or whatever's more in fashion, then." Further, "alone up on this mountain," Julien would cease encountering the "awful ignorance" of "high society." But this mountain is not high enough. In both cases, a better lover of conquest continues his imaginings rather than seize real objects.

Can Julien achieve the destiny of Napoleon? Perhaps in his inward struggles. Wondering while wandering if he can soar as high, he also falls melancholically that he will not achieve greatly. This oscillation is reasonable given his character. Julien mountain hikes again, but feels no peace thanks to those proposed business plans. Julien, "like Hercules," feels "poised in the balance—not between good and evil, but between the mediocrity of assured livelihood and all his heroic, youthful dreams." An assured livelihood is mediocre compared to youthful heroic dreams. Beyond good and evil, he chooses between bourgeois fortune or dreams of grandeur. Julius Caesar wept when at age 32, Alexander had ruled supreme, but he had not yet accomplished as much.<sup>44</sup> Julien weeps as well: "I have no real steadiness," which "makes doubt all the

---

43 Stendhal, Part 1, Chapter 10, 60-61.

44 "We are told that, as he was crossing the Alps and passing by a barbarian village which had very few inhabitants and was a sorry sight, his companions asked with mirth and laughter, "Can it be that here too there are ambitious strifes for office, struggles for primacy, and mutual jealousies of powerful men?" Whereupon Caesar said to them in all seriousness, "I would rather be first here than second at Rome." In like manner we are told again that, in Spain, when he was at leisure and was reading from the history of Alexander, he was lost in thought for a long time, and then burst into tears. His friends were astonished, and asked the reason for his tears. "Do you not think," said he, "it is matter for sorrow that while Alexander, at my age, was already king of so many peoples, I have as

more dangerous.” For “not made of the wood from which great men can be carved, so I worry that eight years spent earning my bread might deprive me of the sublime energy which makes extraordinary things possible.”<sup>45</sup>

Napoleon represents new possibilities in the democratic age for bourgeois Frenchmen. But even with education and career, who can equal his achievements? As Fulcoz says, France stood her tallest in the eyes of the world in his reign of thirteen years: “Back then, there was something grand about everything we did.” But this nostalgia now corrupts the vain passions of equal citizens. Only a historical few extraordinaires can match Napoleon. Can only such few souls be happy in a democracy? Can they even appear in a democracy? Julien is extraordinary, but in spirit. His wars are within. He is not meant for sainthood nor lawgiving, but artistry through his inner personality. Since Rousseau cannot match Napoleon in battle, he provides new frontiers in educations of love. So Julien adapts his stratagems, as he climbs the social ladder, to the life of the mind. His quarrels make for good literature, but they would not make history. He conquers his world, not the world.

---

Julien is one of the few spiritually superior to the many with his Napoleonic love affairs. As his feelings encounter reality, his heroism is his willed hypocrisy, which in Rousseauian fashion contradicts social conventions. *Physis* is his heart, and *nomos* must bend towards it. It is his duty to himself. In this regard, Julien has unswervingly principled hypocrisy in fulfilling his Napoleonic complex. One summer night while speaking, “gesturing, he happened to touch Mme. de Rênal’s hand,” an accidental brush which lites the fuse of dynamite. Julien quickly withdraws his hand, but “it seemed his duty to arrange matters so that this hand would not be drawn back when he touched it.” He did not intend to start a love affair, but it was his duty of superior hypocrisy. “This notion of duty” Julien “was required to accomplish,” alongside his fear of ridicule to feel inferior if he did not finish his mission, “instantly drove all pleasure from his heart” when before enjoying himself.<sup>46</sup>

With women, Julien struggles between his sense of duty and his nervous attitude. In the Great Battle of the Hands with Mme. de Rênal, Julien has “too much contempt for himself and for others not to be aware of the state of his soul.” As a commander, he is still too ignorant, and nervous: “the frightful battle he had to wage against timidity was too far painful for him to notice anything outside himself.”<sup>47</sup> It comes to love or death. Furious at his cowardice to not dare yet make a move, Julien swears, “at the stroke of ten, either I will do what I had been promising to myself all day, or I’ll go upstairs and

yet achieved no brilliant success?”

Plutarch, “The Life of Julius Caesar,” *The Parallel Lives of Plutarch*, Loeb Classical Library edition, Vol. VII, ed. Bernadotte Perrin (1919), 11.3-6, 469, [online](#).

45 Stendhal, *ibid*, Part 1, Chapter 12, 70-71.

46 Stendhal, *ibid*, Part 1, Chapter 8, 50.

47 Stendhal, *ibid*, Part 1, Chapter 9, 51.

blow out my brains.”<sup>48</sup> Keeping her hand in his, “happiness flooded across his heart,” not because he loved her, but the “horrible torture was over.” In his mind, Julien “has done *his duty, performed his act*,” so happy enough to read exploits of Napoleon, yet he feels his duty is, saying “casually” on way to lunch, “I have to tell this woman I love her.”<sup>49</sup>

Thus far, his pride has won over his timidity. And these handholding skirmishes have tired our hero. Warrior Julien needs his sleep. While Mme. de Rênal fell asleep from her happiness, Julien was “mortally exhausted by the struggle in his heart, all day long, between shyness and pride.” Pride reveals all things, as Julien confesses his love of Napoleon to his first lover. But his internal triumph has an external effect. As their love affair continues which they consummate, Julien thinks he deserves better. He leaves her to go back to his room, having “thought it better suited his dignity to leave her in broad daylight, with an infinite recklessness.”<sup>50</sup> His pride triumphs, even at risk of exposure, which Mlle. Elisa enacts. After seminary, Julien meets his match in Mlle. Mathilde. He worships Napoleon, and she her ancestor. Both pridefully play hard to get out of their vanity. Their war games shed light on how democratic pride uplifts and deforms the passions.

Mlle. Mathilde is an example alongside Julien of passion distorted by pride. She demands proofs: “This young girl of high society let her heart be moved when, for good reasons, it ought to be moved.”<sup>51</sup> Now, love as a process can be scientifically understood, according to Stendhal. But to demand logic of what the heart wants is to misunderstand crystallization. Her love is born from her reason. One night, her enrapture by music had Mlle. Mathilde thinking of Julien as Mme. de Rênal had done before. It was love mentally constructed, but not by sentiment: “Mind-made love is of course subtler than true love, but its moments of enthusiasm are limited: it understands itself too well; it is always passing judgement. Rather than deranging the mind, it throbs only to the beating of thought.”<sup>52</sup> Mme. de Renal thinks of Julien by the heart, but Mlle. Mathilde by the head. This contrast suggests why Julien reverses his love from his second beloved to his first.

Public opinion still has its sway. Thanks to lessons from Prince Korasoff, Julien wins the *jeu d'amour*, as Mathilde later carries his child. The Marquis relents and gives Julien an aristocratic title and makes him a military officer. He shall successfully wed into the aristocracy. His inner wars and outer advancement have succeeded. The novel could end here, as Julien moves from one beloved to another, from the black to the red.

48 This quote is missing from p. 51, at the end of the fifth full paragraph, in the Modern Library Edition. Stendhal, Part 1, Chapter Nine, *The Red and the Black: A Chronicle of 1830*, trans. Horace B. Samuel (New York: E.P. Dutton and Co., 1916), [online](#).

49 Stendhal, *The Red and the Black*, trans. Burton Raffel, *ibid*, Part 1, Chapter 9, 52-53.

50 Stendhal, *ibid*, Part 1, Chapter 9, 53; *ibid*, 56-57; Part 1, Chapter 16, 83.

51 Stendhal, *ibid*, Part 2, Chapter 30, 404.

52 Stendhal, *ibid*, Part 2, Chapter 19, 341.

In fact, the Marquis' initial acquiescence mimics the earlier aristocratic desperation of the Chevalier de Beauvoisis. The misunderstanding for the duel, that he is not the man Julien saw at the café, illustrates how interchangeable these social types are. But the Chevalier must by honor duel an equal. In an aristocracy, aristocrats are equals. Here, this class straddles into democracy. Comically, Julien and the Chevalier think over in respectful but awkward talk how to find grounds of honor to duel. As honor can be bought, it can also be faked.<sup>53</sup>

The guillotine, however, cannot be bought. How Julien reacts to the letter by Mme. de Rênal exemplifies democratic love. Immediately riding on his horse to shoot her in church, Julien before forgot entirely Mme. de Rênal when he pursued Mathilde, yet, now the reverse occurs. He falls out of love with Mathilde, forgetting his betrothed and mother to his son, to be back in love with Mme. de Rênal. His crime, seemingly foolish, makes sense not as a love born in the head, but in the heart: such passion throbs beyond the beating of thought, it deranges the mind. But is Julien deranged in *desiring* death? He attempted to kill, and demands to be killed immediately: "Article 1342 of the Code is clear," he tells the judge, "I deserve a sentence of death, and I expect it," for "I'm making myself as guilty as you could ever want," without difficulty nabbing the chased victim. Julien wants not out of jail but life: "Spare me, please, having to deal with you." This is madness.

But does it have method? Used to misery, is death that bad? As Stendhal records, "His whole had been nothing but a long prelude to misery, nor did he have to worry that he might forget what commonly passed as the greatest event of all."<sup>54</sup> An inversed Boethius, Julien takes the perspective of a "philosophical melancholy," as he explains, "I never learned the art of enjoying life until I could see its end closing in on me." He even refuses the hope to escape which Mathilde and Fouqué provide. Despondent, he rebukes his lawyer: "What do other people mean to me? My relationships to *other people* are going to be abruptly cut off."<sup>55</sup> His death sentence is a suicide. He chooses "this greatest event of all." It is not crazy, but deliberate. Why? His remarks feel that, Stendhal comments, thanks to "no exercise," Julien gained "the lofty-minded, wan nature of a young German student."<sup>56</sup> Maybe, Julien Sorel has joined the ranks of bankers and philosophers.

But has his pride entirely left Julien? No. Remorseless, he plans "to scatter pieces of gold among the crowd, as I go to the gallows. Thus joined to the idea of *gold*, my memory would be, for them, a truly resplendent thing."<sup>57</sup> Still Napoleonic, he chooses a death for fame, not with the bourgeoisie or aristocrats, but with the people instead. True, His melancholic philosophy partly results from learning Mme. de Rênal

53 Stendhal, *ibid*, Part 2, Chapter 6, 254-260.

54 Stendhal, *ibid*, Part 2, Chapter 36, 433-434.

55 Stendhal, *ibid*, Part 2, Chapter 39, 452; Part 2, Chapter 40, 454.

56 Stendhal, *ibid*, Part 2, Chapter 44, 477.

57 Stendhal, *ibid*, Part 2, Chapter 36, 435.

survived her attempted murder. Yet he constantly thinks of her, even having Mathilde give away her child to Mme. de Rênal: "In fifteen years, Madame de Rênal will adore my son, and you will have forgotten him."<sup>58</sup> This bloodless command comes from his pride and love. His attempted murder was due to pride and love, as is his death wish and his desire to be remembered by his true peers, the proletariat he defended at his trial, those who valorize Napoleon. And his new desire for forgiveness is also from these vain passions democracy mixes.

Still, the visit of Mme. de Rênal to the prison is ambiguous. She is pious, having repented of her adultery and exposing Julien Sorel. She sincerely believes in God, but "I believe, just as fervently, that the sin I'm committing is frightful." She recalls he fired two pistol shots at her, yet they kiss. She confesses: "As soon as I see you, my whole sense of duty disappears, all that I am is my love for you—though the word *love* is much too feeble. What I feel for you is what I ought to feel only for God: a mixture of respect, and love, and obedience." If Julien told her to attack the jailer with a knife, she would do it before she would know it. Julien has the power to command his beloved to sin. She worships him. He is equal with God to her, if not an idol replacement. She is maybe sacrificing her religion, at least her marriage by her visits. Perhaps Julien is affected by this charity. Love conquers all, says Virgil.<sup>59</sup> Could love for another person enlarge his heart:

Her voice was so sad that Julien embraced her with a happiness utterly new to him. This was not love's drunkenness, but extraordinary gratitude. He had just seen, for the first time, the full extent of the sacrifice she'd made for him.<sup>60</sup>

---

This transformation can be interpreted variously. The pessimistic interpretation is that Julien has his vanity fulfilled. His inner battle is won. Having death means he has won at life. His vision of his identity is now complete. His apotheosis at the end, as Mathilde

<sup>58</sup> Stendhal, *ibid*, Part 2, Chapter 39, 453.

<sup>59</sup> "I will be gone, and the strains I composed in Chalcidian verse I will play on a Sicilian shepherd's pipe. Well I know that in the woods, amid wild beasts' dens, it is better to suffer and carve my love on the young trees. They will grow, and you, my love, will grow with them. Meanwhile, I will roam with the Nymphs on Maenalus, or hunt fierce boars. No frosts will stay me from surrounding with my hounds the glades of Parthenius. Already I see myself traversing rocks and echoing groves; it is a joy to shoot the Cretan shaft from my Parthian bow! Once more Hamadryads and even songs have lost their charms for me; once more farewell, even ye woods! No toils of ours can change that god, not though amide the keenest frosts we drink the Hebrus and brave the Thracian snows and wintry sleet, not though, when the dying bark withers on the lofty elm, we drive to and fro the Ethiopians' sheep beneath the star of Cancer! Love conquers all; let us, too, yield to Love!"

Virgil, *Eclogue* X, Lines 50-69, in Virgil, *Eclogues, Georgics, Aeneid*, trans. H.R. Fairclough, Loeb Classical Library Volumes 63 & 64 (Cambridge, MA: Harvard University Press, 1916), [online](#).

<sup>60</sup> Stendhal, *ibid*, Part 2, Chapter 43, 469-470.

reenacts the legend of Queen Margot kissing the forehead of her beloved Joseph Boniface de La Mole, and then erects a shrine at his tomb, suggests his godlike status, as Mme. de Rênal suggested by her love. And her death, possibly from a broken heart, only three days later proves his conquest over both women. And he grows with his melancholic philosophy to see clearly about the nature of love. As he tells Mathilde, "You surely agree, my dearest, that passions are accidental, in these lives of ours, but they are accidents that happen to superior souls."<sup>61</sup> Passions like love are accidents. Hope dies, faith fails, love ends. But love the passion exists only in superior souls, and Julien Sorel has felt it.

The lessons to learn from this interpretation of literature apply to political science, that of the danger of pride in democracy when not properly channeled.<sup>62</sup> It upsets normal affections, and must be given its proper due, so that the loves of life can be molded into virtuous habits of the heart. But there is a more hopeful lesson, that at death Julien is freed from his worldly pride. His recognition of the sacrifice his beloved makes changes not just how he feels, but how he is. His ability to love is renewed by being loved. René Girard gives this parable: "The conclusions of all the novels are reminiscent of an oriental tale in which the hero is clinging by his finger-tips to the edge of a cliff; exhausted, the hero finally lets himself fall into the abyss. He expects to smash against the rocks below but instead he is supported by the air: the law of gravity is annulled."<sup>63</sup> In this hopeful interpretation,<sup>64</sup> Julien has an apotheosis closer to the books he rejected yet knew well.

61 Stendhal, *ibid*, Part 2, Chapter 39, 452.

62 "That deeper tragedy arises from democratization's influence on pride and self-esteem. As we have seen, the democratization of honors disrupts the social regulation of self-esteem and leaves extraordinary youths to fend for themselves in fashioning self-worth. With the erosion of social or collective standards, those young people begin to feel a new need to authorize or even author their own standards of distinction. Historically speaking, this need of pride is a new feature and effect of the democratic age. Born of that new need, Julien's tragedy foretells the tragedy of democracy's children—Stendhal's prophecy for liberal democracy. Indeed, as we may wish to discuss, the more numerous the rights our regime assigns to us, the more sharply we may feel the self-deifying imperative of modern pride. This is not to say that from Stendhal's historical point of view, the youths of the future are fated to love as Julien and Mathilde do, though love may be a likely servicer of pride's needs. Foremost, Stendhal's prophecy is that pride's new need will subjugate our wills to the rule of accidents. More than merely casualties of contingency, we will become the willing slaves of contingency. ... Julien does not grasp what Stendhal expects future readers of *The Red and the Black* to consider: namely, that his protagonist's tragedy is a distinctly historical tragedy, and that even if Julien's tragedy of contingent loves is now our tragedy, there was at least a point in the past when the development of modern love did not have to veer in our direction."

Jeffrey Smith, "Stendhal's Prophecy for Liberal Democracy," *ibid*.

63 René Girard, *Deceit, Desire, & the Novel: Self and Other in Literary Structure*, trans. Yvonne Freccero (Baltimore: The Johns Hopkins University Press, 1976), 294.



But why is there only a happy few in a democracy? “To the Happy Few” was a phrase Stendhal frequented.<sup>65</sup> That dedication alludes to *The History of Henry V*, Act 4, Scene 3, where Shakespeare has King Henry give his “St. Crispin’s Day” Speech before his men at the Battle of Agincourt (1415). There he proclaims with aristocratic gusto, “If it be a sin to covet honor, / I am the most offending soul alive.” He further exclaims his ennoblement of the few men who fight for their conquering king upon this day, unlike the majority of gentlemen asleep in England: “This story shall the good man teach his son; / And Crispin Crispian shall ne’er go by, / From this day to the ending of the world, / But

64 “Repudiation of the mediator implies renunciation of divinity, and this means renouncing pride. The physical diminution of the hero both expresses and conceals the defeat of pride. Once sentence with a double meaning in *The Red and the Black* expresses beautifully the link between death and liberation, between the guillotine and the break with the mediator: “What do *Others* matter to me,” exclaims Julien Sorel, “my relations with others are going to be abruptly cut off.”

“In renouncing divinity the hero renounces slavery. Every level of his existence is inverted, all the effects of metaphysical desire are replaced by contrary effects. Deception gives way to truth, anguish to remembrance, agitation to repose, hatred to love, humiliation to humility, mediated desire to autonomy, deviated transcendence to vertical transcendence.

“This time it is not a false but a genuine conversion. The hero triumphs in defeat; he triumphs because he is at the end of his resources; for the first time he has to look his despair and his nothingness in the face. But this look which he had dreaded, which is the death of pride, is his salvation. ... Julien Sorel rejects Others and embraces solitude...Julien wins solitude but he triumphs over isolation. His happiness with Mme de Rênal is the supreme expression of a profound change in his relationship with Others. When the hero finds himself surrounded by a crowd at the beginning of his trial, he is surprised to find that he no longer feels his old hatred for Others. He wonders whether Others are as bad as he once thought them. When he no longer envies people, when he no longer wishes to seduce or dominate them, then Julien no longer hates them. ... Stendhal attributes Julien Sorel’s “German mysticism” to the extreme dampness of his prison cell. But the conclusion of *The Red and the Black* remains a meditation on Christian themes and symbols. In it the novelist reaffirms his skepticism but the themes and symbols are nonetheless present in order to be clothed in negations....We shall see everything which touches on these themes, including the monastic vocation which of Stendhal’s heroes, in a fresh light which the author’s irony cannot hide from us.”

René Girard, *Deceit, Desire, & the Novel*, *ibid*, 294-295, 311.

65 “At the end of three of his books – *Promenades dans Rome* and his two masterpieces, *Le Rouge et le Noir* and *La Chartreuse de Parme* – Stendhal (quoting from Shakespeare’s *Henry V*, Act 4, Scene 3) inscribed the same words, in English: TO THE HAPPY FEW. He often expressed the belief that he would find readers after a hundred years (he could not have known how right he was!). Meanwhile his readers were dismally few – a situation that once provoked the jibe from his ‘bookseller’: “Your books are truly sacred: no one touches them!””

Simon Leys, *With Stendhal*, introduction, translation, and notes by Simon Leys (Melbourne, Australia: Black Inc. 2010), Endnote 15, 75-76.

we in it shall be remember'd. / We few, we happy few, we band of brothers; / For he to-day that sheds his blood with me / Shall be my brother; be he ne'er so vile/ This day shall gentle his condition." King Henry enlarges nobility to those who fight in war.

In that battle, the few were happy. The French have 60,000 men, the English have 12,000, yet the happy few win gloriously. The English longbow had defeated the French cavalry. If Julien indicates the tendency of democratic passions, then few are happier in democracy than aristocracy, perhaps. For Stendhal, there is some reprieve of spontaneous passion in Fabrizio del Dongo, an Italian aristocrat. In his travel memoirs, Stendhal identifies Italy with aristocratic enchantment.<sup>66</sup> The democratic levelling, like the longbow, removes aristocratic hierarchies that properly channel pride. Even where Tocquevillean democracy seems flourishing, there were losses. In America, as in France, public opinion intrudes into the private lives of democratic souls. Yet de Tocqueville knows that a democratic nation needs to emphasize inherited aristocratic elements to chasten its citizens. This premodern moral capital finances the republican principles for self-rule to flourish.

Stendhal creates fictions to protect private life. Yet, he says in *The Charterhouse of Parma*—"Where I am going,' I told my friends, 'there will be no parties like this one, and to while away the long evenings I shall make a novel out of your story'." Unlike life at the Canon's Salon, where Stendhal parties with the nephew and learns a story about the Duchess Sanseverina, no such goods happen in France. Public opinion reigns in France, England, and the United States, but in Italy? The contrasts between France and Italy are stark. Stendhal must be careful, for "this tale is anything but moral," and "the French" take "pride" in "gospel purity." The dichotomy within France is French pride in gospel purity versus "the reputation of an assassin." To speak of Italian private life in a story wins ill repute. In Italy, things are different: "hearts in that country differ altogether from those in France." Public opinion has not yet invaded. Democracy has not yet arrived.

66 "There, seated upon the step of a faldstool, with my head thrown to rest upon the desk, so that I might let my gaze dwell on the ceiling, I underwent, through the medium of Volterrano's *Sybils*, the profoundest experience of ecstasy that, as far as I am aware, I ever encountered through the painter's art. My soul, affected by the very notion of being in Florence, and just by the proximity of those great men whose tombs I had just beheld, was already in a state of a trance. Absorbed in the contemplation of *sublime beauty*, I could perceive its very essence close at hand; I could, as it were, feel the stuff of it beneath my fingertips. I had attained to that supreme degree of sensibility where the *divine imitations* of art merge with the impassioned sensuality of emotion. As I emerged from the porch of *Santa Croce*, I was seized with a fierce palpitation of the heart (that same symptom, which, in Berlin, is referred to as an *attack of nerves*); the well-spring of life was dried up within me, and I walked in constant fear of falling to the ground."

Stendhal, "A Journey from Milan to Reggio," *Rome, Naples, and Florence*, trans. Richard N. Coe (London: Calder Publications, 1959, 2010), 302.

The dichotomy between French hearts and Italian hearts is stark. Italians, “sincere, honest people,” when “not intimidated will say what they think,” and “only intermittently are they subject to vanity,” which on such occasion “becomes a passion” by name of “*puntiglio*,” but this passion does “not hold poverty up to ridicule.” The Italians are not tainted by love of commercial interests that coincide with democratic peoples. The gospel pure French are opposite, without the Italian “irregularities of nature,” having (sarcastically) “lofty morality and the graces” who “love money above all things and never sin out of love or hate,” while “Italians in this tale are virtually opposite.” The French public is in love with financial gain, while the sincere, occasionally vain, Italians sin out of love and hate, passions, that is. This dichotomy of nations extends further to regimes.

It is Italy versus England, France, and America, where public opinion reigns. In Italy, private life still subsists. Stendhal reinforces this point: “each time we venture two hundred leagues from South to North, we confront a new novel as well as a new landscape.” Here in Parma is seen the alternative to *The Red and the Black*, a different novel and landscape where “characters, seduced by passions” Stendhal cannot share, happen to “descend to profoundly moral actions,” for “such things are no longer done in a country where the sole passion surviving all the rest is for money, the means of vanity” (105). So is *puntiglio* aided by passion for the money? Is private life here also eaten up by monetary effects of public opinion reigning on heartstrings and purse strings? This issue of private life versus public opinion, passion versus money, consolidates into the ruin of aristocratic, old world honor privately held in conversational esteem by democratic levelling.<sup>67</sup> Does democracy gone wrong beat and bloody men of spirit until they turn black and red all over?

67 Stendhal, *The Charterhouse of Parma*, translated by Richard Howard, illustrations by Robert Andrew Parker (New York: The Modern Library, 2000), xv, xiv, 105.

# The Righteous Mind by Jonathan Haidt

## Part I: Intuitions Come First, Strategic Reasoning Second

*“The mind is divided, like a rider on an elephant, and the rider’s job is to serve the elephant.”*

### Chapter 1: Where Does Morality Come From?

Different cultures include different concerns under the umbrella of morality. Take the following story: *A man goes to the supermarket once a week and buys a frozen chicken. Before cooking the chicken, he has sex with it. Then he cooks it and eats it. In this story, there is no harm to the chicken.* When Jonathan Haidt asks well-educated people studying at universities in Western countries, they often feel a flash of disgust, but hesitate to say the family had done anything *morally* wrong. This is especially true of politically liberal Westerners. But when Haidt asks people from non-Western countries, which comprises most of the world, people find the chicken story morally wrong. This shows that for different groups of people, morality can encompass a wider range of behaviors than just those actions that harm another person or being.

Our sense of morality begins developing as kids. Elliot Turiel studied the development of moral views in kids by telling them stories about other kids breaking rules. He’d then ask if what the kid in the story did was okay. Some of the stories go as follows: *A boy goes to school that requires uniforms. One day he doesn’t wear his uniform. Did he do something wrong? What if his teacher said it was okay?* Turiel found that kids as young as five usually say it is wrong for the boy to break the rule, but okay if the teacher gave him permission. That is, kids recognize that rules about many things in life (like clothing, food, etc.) are *social conventions*—arbitrary and changeable.

But kids recognize that some rules are not social conventions. If you ask kids about actions that hurt other people (for example, tell them a story about a girl pushing a boy off a swing) nearly all kids say the girl was wrong. Even if you tell them that the teacher said it is okay, they’ll still say the action was wrong. This shows that kids recognize rules which prevent harm are moral rules—special, important, unalterable, and universal. Studies like these led Western psychologists to conclude that harm was the basis of all morality.

This was wrong—or at least, it was only correct in (some) Western cultures. Other cultures have a far wider range of actions and behaviors that, to them, are part of the moral matrix. For example, the Azande believe in witches (which can be men or women). The Azande are careful not to make neighbors angry or envious, because that will get them pegged as a witch by neighbors. Their belief in witches, tied as it is to asocial behavior, shows morality helps order human society.

Another example is the Ilongot, a society where young men can gain honor by cutting off people’s heads. Some killings are made in revenge, but many are against strangers who the murderers never knew. Small groups of men channel resentment and in-group frictions through a group-strengthening “hunting party” that, after the killing is done,

ends in a long night of communal celebratory singing. The Ilongot's use of honor killings shows how morality helps address and ease tension within the group and competition between groups.

A third example is the Hua of New Guinea. The Hua have an elaborate network of food taboos that govern what men and women can eat. Men can't eat any food resembling a vagina if they want to become men. The Hua know these are social conventions—they don't believe other tribes have to follow them. This is an example of how taboos can govern duties and relationships to others in your society.

The idea here is that if morality is just about harm, then it makes no sense that most non-Western cultures moralize about so many practices—from sex, to food, witchcraft—that have nothing to do with harm. In fact, *even in the West*, not all subcultures limit morality to the domain of harm. Liberals tease religious conservatives for being sexually prude, given their taboos over non-missionary sex positions; meanwhile, conservatives tease liberals for having ample taboos about food—be that fair-trade coffee, Paleo, or whatever the new fad toxin one should avoid.

Richard Shweder argued that these moral strictures come from a broader human need than just avoiding harm—humans need to figure out how to *organize society*.

Specifically, they need to sort out how to balance the needs of individuals versus groups. Most societies have chosen a sociocentric answer. They put the needs of the group and institutions first, and individuals' needs second. Other societies have chosen an individualistic approach. The needs of individuals come first; society serves the individual.

The sociocentric approach dominated most of the ancient world, but the individualistic answer became a powerful rival during the Enlightenment era in Europe, and has since surpassed the sociocentric approach since the 1900s.

To show how society's decision between sociocentric versus individualistic moral culture can lead to wildly different moral views, Richard Shweder conducted a study on Indians and Americans. He asked 180 Americans (kids and adults) and 180 Indians (kids and adults) whether a series of scenarios were morally right or wrong. Stories would go as follows: (A) *While walking, a man saw a sleeping dog. He approaches it and kicks it.* (B) *A young married woman goes alone to see a movie without telling her husband. When she returns, he said if she does it again he will beat her. She does it again, and he beats her.* (C) *A widow in your community eats fish two to three times per week.*

Schweder came away from this study with two interesting findings. First, kids answered almost exactly as adults of their own culture did. Second, Americans and Indians did not agree on many scenarios. While vast majorities of both groups agreed it was morally wrong for the man to kick the dog, Americans thought the husband was morally wrong to beat his wife, while Indians thought it was okay. Meanwhile, Indians thought the widow was morally wrong to eat so much fish, while Americans thought it was okay.

Interestingly, the Americans even went further, often arguing that it was morally wrong to tell the widow she *couldn't* have fish. This implies that when Americans distinguish between what is a universal moral rule and what is a mere social convention, they take an individualistic approach. When a rule binds the individual unfairly, it is an immoral rule. But people from other studies don't share this view. Turiel's studies on

Western kids distinguishing between harm and social convention were misguided because they only looked at the morality of Western cultures. Broadening the sampled cultures revealed that morality encompasses much more than the norm against harming others.

Haidt did his own study composed of 360 interviews of participants from Porto Alegre, Philadelphia, and Recife. The participants in each city were either high class or low class, and either children or adults. He read them stories like the “chicken sex story,” then asked if there was anything morally wrong with the action, taking care to point out that no one was harmed. Most participants felt that violating harmless was universally wrong. But there were some nuances. People from Philadelphia were more likely to distinguish between moral rules and social conventions than were people from Porto Alegre and Recife. Upper-class people in Philadelphia judged the harmless taboos as violation of social conventions; lower-class people from Recife found them to be violations of universal moral rules. Porto Alegreans were mixed. Social class turned out to be a stronger predictor than geographic location of whether people saw the taboos as social conventions or moral violations.

Haidt also observed that participants condemned the action instantly, then took time to try to invent victims for the stories. This was despite that Haidt had written the stories to carefully remove all conceivable harm to others. About 38% of participants tried to find someone who was harmed. This was true of participants across all regions and ages. Participants would often offer a possible victim half-heartedly, almost apologetically. Many of these possible victims were implausible—one child trying to reason why flushing the American flag down the toilet argued that the clutter might clog the toilet, causing it to overflow, which would require an adult to fix it. Even after the interviewer gently corrected the participant, pointing out there was no victim in the story, they would still keep searching for another victim. They’d often say things like, ‘I know it’s wrong, but I just can’t think of a reason why.’

This was Haidt’s first big piece of evidence that *people use moral reasoning to justify moral emotions*.

## Chapter 2: The Intuitive Dog and its Rational Tail

“When a group of people make something sacred, the members of the cult lose the ability to think clearly about it. *Morality binds and blinds*. The true believers produce pious fantasies that don’t match reality, and at some point somebody comes along to knock the idol off its pedestal.” So explains Haidt the phenomenon of moral sanctification by human tribes.

Moral feelings are different than moral reasoning. Antonio Damasio studied patients who had damage to their ventromedial prefrontal cortex (vmPFC). They had no emotionality. They could look at joyous or gruesome photographs and feel nothing. They still retained their full IQ, though, and could distinguish between morally right and morally wrong actions. They just couldn’t make decisions in their personal lives and at work. Their lives fell apart.

Damasio concluded that gut feelings and bodily reactions were necessary to think rationally. To these patients, every option *felt* equally good, so they couldn’t make decisions without weighing every pro and con with conscious, verbal reasoning.

However, human short-term memory quickly becomes over-taxed once options reach 6 or 7 in number. Emotions help us choose when options get complicated. Hence why the patients couldn't make decisions—they *needed* emotional passions in order to function properly.

Haidt did his own study to test if people's moral judgments used the rational part of the brain or the emotional part. He would ask participants to hold either (a) a string of numbers, such as 7250475, in their mind, or (b) just one number, such as 7. He then asked them to simultaneously perform another task. If they did worse at the second task when trying to remember 7250475 than when they were trying to remember 7, Haidt concluded that they were using their rational brain for the second task. Hence why they did worse with the long string of numbers.

Interestingly, it turns out that people were able to answer moral questions when they were remembering long strings of numbers just as easily as when they were trying to remember one digit. This suggests that people don't use rational thinking to answer moral questions—they use moral emotions and feelings.

Haidt investigated this further in a study with Scott Murphy. They interviewed 30 subjects on their reactions to several scenarios. One scenario asked whether people would eat a sterilized cockroach. (This pits rational thinking against our sense of purity—the roach is sterile (harmless) but gross (impure)). A third scenario asked if participants would sign a piece of paper selling their soul for \$2. Below the signature line, though, was written, "This form is part of a psychology experiment. It is NOT a legal or binding contract in any way."

Only 37% would eat the cockroach and only 17% would sign the paper. Murphy asked the participants who refused why they wouldn't do so. He tried to convince them to do it. Still, most people clung to their original positions, even though some who refused to sign their souls away were atheists! Haidt summarizes the findings thusly: "People made moral judgments quickly and emotionally. Moral reasoning was mostly just a post hoc search for reasons to justify the judgments people had already made."

People are very good at offering justifications for their immediate reactions to moral scenarios. Howard Margolis performed a study that asked people to solve a simple task. He laid out four cards. The first card had an E on it, the second K, the third a 4, and the fourth a 7. The task was to choose the smallest number of cards that you must turn over to decide whether the following rule was true: "If there is a vowel on one side, then there is an even number on the other side." The answer is two—the E and the 7 cards.

But most people thought they had to turn over the E and the 4 cards. They didn't realize that turning over the 4 card could produce a finding that would not invalidate the rule. When Margolis asked people why they chose their answer, mostly everyone was able to offer an explanation. And mostly everyone was just as confident in their reasoning whether they got the answer right or wrong. Margolis concluded that judgment and justification are two separate mental processes. In his words, "Given judgments, human beings produce rationales they believe account for their judgments. But the rationales are only ex-post rationalizations."

Margolis called the two processes "Seeing That" and "Reasoning Why." "Seeing That" is the process of pattern matching. It is rapid, automatic, and effortless. It drives our perceptions of real world objects. All animals do this as they move around the Earth. "Reasoning Why" is our human brain's process of describing how we think once we've

reached a judgment. This is only seen in animals with language and who thus have a need to explain themselves to others. It is conscious, not automatic, and is easily disrupted when the brain is also working on other cognitive tasks.

Margolis' theory easily explains the Haidt's findings. People "see" that eating a cockroach is the wrong pattern or behavior, but can't easily justify why those are the right actions after the fact. Haidt concludes, "We do moral reasoning not to reconstruct the actual reasons why we ourselves came to a judgment; we reason to find the best possible reasons why *somebody else ought to join us* in our judgments."

Underlying these moral intuitions is a broader phenomenon that applies to all our emotions. When we form emotional reactions, we begin by appraising something that just happened to us based on whether it advanced or hindered our goals. These appraisals are a kind of information processing. Emotions aren't dumb—Damasio's patients made *worse* decisions when they were lacking emotions. Moral judgment, then, is as much a rational process as conscious thinking, albeit an automatic one.

These automatic processes are the basic hardware of the brain—the conscious thinking is just an add-on. The human mind evolved from the same software that ran animal brains for 500 million years. When human brains evolved language 1 million years ago, the brain did not reinvent itself. Rather, the new language-based part of the brain just added itself on to the rest of the brain that already existed.

Haidt calls our language-processing part of the brain the "rider." He calls the ancient part of the brain the "elephant." The rider can do several useful things to serve the elephant. It can see further into the future by considering alternative possible scenarios and choosing the optimal one. It can learn new skills and master new technologies to help the elephant reach its goals and avoid disasters. It can be a spokesperson for the elephant, coming up with post-hoc explanations defending whatever the elephant just did.

Post-hoc justifications are, by the way, extremely valuable. Once humans created language, they used it to gossip about each other. Gossip prevents free-riding and amoral acts by damaging the bad actor's reputation when he's not around. He pays the price for a bad act today through others' disapproval later. Given this, it became extremely valuable for individuals to be able to quickly justify their actions to others. As Haidt describes it, the elephant needs a "full-time public relations firm." The part of the brain that processes language thus also keeps itself busy managing your reputation, building alliances, and recruiting bystanders to support your side in our common daily disputes.

This process means, however, that humans are dreadful at seeking out evidence that might disconfirm our initial judgments. Sometimes friends can help us see our blind spots or appreciate different perspectives. They can challenge our viewpoints, give us reasons and arguments that trigger new intuitions, and make it possible that we change our minds on things. But even friends rarely change an individual's mind. Rather, the most common cause for people to change their mind is when we desire to join the opinion of someone popular.

If moral and political arguments seem frustrating, it is because moral reasoning is just a made-up justification for our emotional intuitions. If you want to change people's minds, you shouldn't try to reason with their rational rider, but with their intuitive elephant.



## Chapter 3: Elephants Rule

Brains evaluate instantly and constantly. They evaluate everything in terms of “potential threat” or “benefit to the self.” They then adjust behavior to get more good stuff out of life and less bad stuff. To coordinate behavior, the brain stimulates an “affect,” which is a small flash of positive or negative feeling that prepares us to approach or avoid something.

Affective reactions can be simplistic mechanisms. Robert Zajonc showed that people will prefer something simply because they’ve seen it before several times. He showed people arbitrary things like Japanese pictograms, words in a made-up language, and geometric shapes. He asked participants how much they liked these things. People liked those objects and words that they had seen a few times before. This is, you’ll notice, a basic principle of advertising.

Social and political judgments are particularly intuitive. A study looked at people’s ability to determine if the second word in the following pairs was “good” or “bad”:

- Flower—Happiness
- Hate—Sunshine
- Love—Cancer

It took people, on average, 250 milliseconds longer to recognize that ‘sunshine’ was a good word than it took them for ‘happiness,’ simply because sunshine was preceded by ‘hate,’ while ‘happiness’ was preceded by ‘flower.’ That is to say, “hate” primed people’s minds for a negative evaluation, causing their brains to lean that direction. They then needed an extra 250 milliseconds to undo the priming.

Another study found that this same phenomenon applies to political groups. The study replaced the first word in the two-word pairs with a political word. For example, ‘flower’ might be replaced by ‘Clinton’; ‘love’ by ‘pro-life.’ Liberals were able to categorize the good words (like ‘sunshine’) faster than conservatives if the preceding word was ‘Clinton’ and ‘pro-choice.’ The conservatives were better when the preceding word was “Bush” or ‘pro-life.’

This behavior applies to our judgments of social and political groups. The Implicit Association test flashes photographs of people of different races and then asks participants to categorize a word as good or bad. Just like the study above, people take longer to categorize a word like ‘sunshine’ as good after seeing pictures of certain racial groups. The IAT has found that most people have a negative implicit association with many social groups like blacks, immigrants, obese people, and the elderly.

Another study looked at how fast people form impressions of others. Alex Todorov gathered photos of the winners and runners-up in hundreds of US Senate and House of Representative races. He showed participants just the photographs of the two political contenders and asked them to pick who seemed more competent. Two-thirds of the time, people chose the photo of the candidate who actually won! Todorov then ran another experiment where he showed people the photograph for only one-tenth of a second. Amazingly, people were *just* as accurate. Haidt writes, “Whatever the brain is doing, it is doing it instantly.”

Even babies show signs that they come with innate abilities to understand their social world. Psychologists Kiley Hamlin, Karen Wynn, and Paul Bloom put on puppet shows

for 6-month and 11-month babies. One puppet would try to “climb” a hill. Sometimes a second puppet came to “help” the first one up the hill. Other times a third puppet would try to “bash” the puppet down the hill. After the show, babies were presented with the “nice” and “mean” puppets on a tray. Babies consistently reached out to touch the nice puppet. This shows that babies watch how people behave toward other people and develop preferences for those who are nice, not mean.

If we slow down the brain’s immediate emotional reaction to a moral dilemma, sometimes the brain will process more logical arguments. Joe Paxton and Joshua Greene asked college students about the following story: Two siblings, a brother and sister, want to have sex. They use perfect contraception, no one finds out, and no one is harmed. The researchers gave half of participants a “weak” justification for this conduct: “If Julie and Mark make love, then there is more love in the world.” They gave the other half of their participants a “strong” justification: “Aversion to incest is really caused by an ancient evolutionary adaptation for avoiding birth defects in a world without contraception, but because Julie and Mark use contraception, that concern is not relevant.” The subjects in both groups condemned the incest nonetheless, and did so in equal proportions. However, when researchers required some subjects to wait 2 minutes before they declared their judgment, these people were substantially more tolerant toward Julie and Mark’s intercourse.

## Chapter 4: Vote for Me (Here’s Why)

If 400 insects are working toward a common goal, you can bet they are all siblings. When 100 people work together on a construction site or march off to war, you wouldn’t expect them to be part of one family. That’s weird. Human beings are extremely unique in our ability to cooperate with other humans to whom we are not biologically related. We do this by forming formal and informal systems of accountability. As Haidt writes, “We are really good at holding others accountable for their actions, and we’re really skilled at navigating through a world in which others hold us accountable for our own.” Phil Tetlock is a researcher in accountability. He describes us all as “intuitive politicians” striving to maintain “appealing moral identities” in front of multiple “constituencies.” Tetlock has done studies giving people information about a legal case and asked them to infer guilt or innocence. Some subjects are told they’ll have to explain their decision to someone else; others are told they won’t have to explain their decisions. When people know their decisions won’t be reviewed by someone else, they display the usual catalogue of errors, laziness, and reliance on gut feelings that have been documented in decision-making research. When people know they’ll have to explain themselves, though, they think more systematically and critically.

Tetlock found two different kinds of careful reasoning. First, there is Exploratory thought, which is an evenhanded consideration of alternative points of view. Second is Confirmatory thought, which is a one-sided attempt to rationalize a particular point of view. Tetlock believes people will take an exploratory approach only when three conditions hold: 1) Decision makers learn before they form any opinion that they will be accountable to an audience; (2) The audience’s views are unknown, AND (3) The decision maker believes the audience is well-informed and interested in accuracy.

The problem is that people are very good at coming up with reasons to support positions they *want* to be true. When subjects are told that an IQ test gave them a low score, they chose to read articles criticizing the validity of IQ tests. When subjects are told that an IQ test gave them a high score, they chose to read articles supporting the validity of the tests. When people read a fictitious scientific study that reports a link between caffeine and breast cancer, women who are heavy coffee drinkers find more flaws in the study than do men and more than women who drink less caffeine. When participants are asked to lick a strip of paper to determine whether they have a serious enzyme deficiency, they'll wait longer for the paper to change color when the color change is desirable than when the color change indicates a problem. Essentially, the brain is playing one of two games: Either the position is one the brain wants to believe, in which case the brain asks itself, "Is there *any* evidence justifying my belief?" Or, the position is one the brain wants to be false, in which case it asks itself, "Is there any reason I *don't* have to believe this?"

Slightly related to this motivated reasoning, the brain will try to cheat *to the degree it can justify cheating*. To illustrate, researchers performed a series of studies where participants could get more money by solving math problems. The study was designed to allow participants to cheat without getting caught. Researchers found that the majority of people cheated, but only a *little* bit.

Our need to validate our beliefs bleeds right into our political views, but, interestingly, not in the service of the self. Rather, our cognitive biases apply to our *group*. Decades of research on public opinion show that self-interest doesn't predict well which public policies voters prefer. Parents of kids in school are not necessarily more supportive of government aid for schools than other citizens. Young men subject to the draft are not more opposed to military escalation than men too old to be drafted. People without health insurance aren't more likely to support government-issued health insurance than people who are covered.

Rather, what people care about are their racial, regional, religious, or political tribes. Political opinions serve as "badges of membership." Drew Westen used fMRI scanners to watch participants' brains while presenting them statements from either President George Bush or his Democratic challenger, John Kerry. All the participants had described themselves as either strongly Republican or strongly Democratic. When the statements challenged their preferred candidate, the participants' brains showed activity near the brain areas associated with negative emotion and responses to punishment. It was as if the challenge to their group's leader actually caused them physical pain.

Meanwhile, there was no activation of the dlPFC, which is the brain area for thinking coolly rationally. That is to say, people were not evaluating these statements objectively.

Westen finished his study by asking participants to justify why the original, threatening statement wasn't so bad. At this point, the part of the brain associated with *pleasure* activated. This shows that hearing good things about one's candidate give you a hit of happiness.

A running theme in Western thought has been that thinking more logically about ethics will cause people to act better. But expertise in moral reasoning does not seem to improve moral behavior. Eric Schwitzgebel measured how often moral philosophers give to charity, vote, call their mothers, donate blood, donate organs, clean up after

themselves at philosophy conferences, and respond to emails from students. In none of these ways are moral philosophers better than other philosophers or other professors. In fact, Schwitzgebel found that academic books on ethics (presumably borrowed by ethicists) are more likely to be stolen or never returned.

## Part II: There's More to Morality than Harm and Fairness

*"The righteous mind is like a tongue with six taste receptors."*

### Chapter 5: Beyond WEIRD Morality

When Haidt performed his studies on people's reactions to the chicken sex story and other morally queer tales, he found that his University of Pennsylvania students were outliers. Working class Americans and rich and poor Brazilians all reacted differently than Haidt's college students.

Haidt found his UPenn students were uniquely devoted to the harm principal—that the only justification for preventing someone from doing something was to prevent harm, and nothing else. They were the only group that frequently ignored their own feelings of disgust and said an action which bothered them was morally permissible. They were the only group in which a majority (73%) were able to tolerate the chicken story.

In 2010, cultural psychologists Joe Henrich, Steve Heine, and Ar Norenzayan reviewed dozens of studies showing that people from Western, educated, industrialized, rich, democratic (i.e. W.E.I.R.D.) countries were morally peculiar from other human societies. They are the least typical, least representative people to study when learning about human nature. Even within WEIRD cultures of Europe and the US, educated, upper-middle-class people (like Haidt's UPenn students) were the most unusual of all. They are *highly* individualistic.

Citizens of WEIRD societies are more likely to see the world as full of separate objects, rather than relationships. One example of this: Westerners are more likely to finish the sentence, "I am..." with a list of their internal psychological characteristics (e.g. 'happy,' 'outgoing,' etc.). East Asians are more likely to finish with their roles and relationships (e.g. 'husband,' 'son,' etc.). This carries into our perception of the physical world: Westerners are better able to remember the objective length of a line on a piece of paper than are East Asians; East Asians are better at remembering the relative size of a line to a square. Westerners are also more likely to think analytically (i.e. separating an object from its context, assigning it a category, and assuming what's true about the category is true about the object). People from most other societies are better at thinking holistically about an object and its context together.

This all spills over into morality. If you live in a WEIRD culture, you think more about individuals and their individual rights. If you live in a non-WEIRD culture, you think more about the needs of groups and institutions. For holistic societies, a morality based simply on harm and fairness will not be sufficient; people in those societies believe there is far more to morality than harm or fairness.

Getting back to Richard Shweder, he analyzed 600 interviews on morality between Indians and Americans to understand the moral foundations of each culture. He

concluded that Western culture is based largely on the ethic of *autonomy*: people are, first and foremost, autonomous individuals with wants, needs, and preferences. People therefore should be free to satisfy these as they see fit. Autonomous societies develop moral concepts like rights, liberty, and justice.

Non-Western cultures, meanwhile, value *community* and *divinity* alongside autonomy. *Community* is based on the idea that people are members of larger entities like families, teams, armies, companies, tribes, and nations. These groups are larger than the sum of their parts; they matter and must be protected. People have assigned roles and obligations to uphold the group. These societies have moral concepts like duty, hierarchy, respect, reputation, and patriotism. Too much individualism (i.e. people pursuing their own personal goals) could weaken the social fabric and destroy the institutions upon which everyone depends. *Divinity* is based on the idea that people are temporary vessels for a divine soul. We should behave accordingly. This is why the chicken sex story, involving no harm and violating no rights, seems so degrading. Too much individualism can become degrading hedonism.

Haidt applies this framework to America. America is made up of sub-groups, each of whom have their own moral combination of autonomy, community, and divinity. For example, Conservatives believe that a red, white, and blue cloth is a “flag” that represents the American community in a special way. Liberals, meanwhile, take the image of Martin Luther King Jr. as sacred, and would probably react strongly to submerging his likeness in a jar of urine, as was done to the Virgin Mary by an artist in NYC. Haidt concludes that many moral matrices coexist even within a single nation. “Each matrix provides a complete, unified, and emotionally compelling worldview, easily justified by observable evidence and nearly impregnable to attack by arguments from outsiders.”

## Chapter 6: Taste Buds of the Righteous Mind

Haidt identifies six different “taste buds” in the human mind. Just like the human tongue can taste sweet, sour, salty, bitter, and savory, humans have a taste bud for Care, Fairness, Loyalty, Authority, and Sanctity.

Similar to taste buds, our cultures can shape what tastes we prefer. But it isn’t “anything goes.” You won’t see a culture basing its entire cuisine on tree bark or bitter taste. Nor will you see a certain culture base its moral matrix on, say, authority alone. Haidt argues that we developed these “moral taste buds” to serve an evolutionary need. For example, our aversion to *harm* and our preference for *care* (remember the babies watching the puppet shows?) evolved to encourage us to protect and care for children. The original evolutionary trigger was suffering, distress, or neediness in one’s own child. Today, however, even cute animals can trigger our desire to care and detestation of harm.

Our obsession with *fairness* and fury at *cheating* come from our need to build cooperative partnerships. The original trigger was cheating and deception by those we trust, maybe during a hunt or during hungry winters. But today, even a broken vending machine can set off our fury at being “cheated.”

*Loyalty* and *betrayal* evolved from our need to form cohesive coalitions. The original trigger was threat to our group from hostile neighboring tribes. Today, sports teams and nation-states trigger the same moral emotions.

*Authority* and *subversion* help us forge beneficial relationships within hierarchies. The original need was to maintain stable hierarchies in tribal bands. Now we do the same within corporations, inventing things like bosses and respected professionals.

Finally, *sanctity* and *degradation* emotions exist to avoid contamination. The original triggers were waste products and diseased people, among others. Today though, even “toxic ideas” like communism and racism trigger our feelings of sanctity and degradation.

## Chapter 7: The Moral Foundations of Politics

### The Care-Harm Foundation

Mammals have few offspring. Each one becomes a high-stake bet that requires a lot of investment. Mammals must care for and nurture their offspring for a long time. Humans, which produce children with such large brains that they must be birthed before even being able to walk, have the fewest children and require pairs to help the mother during pregnancy, delivery, and caring for the child for years. Evolution thus favored men and women who had automatic reactions to the signs need and suffering from their children. (Hence why parents react so quickly to their baby’s crying). These reactive parents were the most likely to successfully raise their fragile children to adulthood.

This is no “just-so” story. It is Attachment Theory, a well-supported theory that describes the system by which mothers and children regulate each other’s behavior so that the child gets a good mix of protection and opportunities for independence.

Both major political tribes in America (conservatives and liberals) have Care woven into their moral matrices. For example, liberals take on concerns like animal welfare and genocide in Darfur. Conservatives, meanwhile, are concerned about the welfare of veterans and offer support through groups like the Wounded Warrior Project. There are differences, though. Liberal Care is often more universal (for all humans, for all species) compared to conservative Care (for those who were injured serving the group loyally).

### The Fairness-Cheating Foundation

Robert Trivers was one of the first researchers of reciprocal altruism. He began by observing that evolution can create altruistic behavior in certain special situations: When individuals could remember their prior interactions with others and limit their niceness to those who were likely to repay the favor. In these rare situations, the individual who survived best was the one who could correctly identify others who were most likely to cooperate without cheating.

Humans do this all the time. Hunters work together to hunt large prey no one can catch alone. Neighbors watch each other’s houses and loan each other tools. People who play tit-for-tat reap benefits from cooperation more than those who never form those relationships with others. To maintain the cooperation, people have moral emotions like

pleasure, affection, and friendliness toward those who can be trusted, and anger, contempt, and disgust at those who cheat.

Today, this manifests itself in many political ways. Liberals distrust groups that are perceived as wealthy and powerful because those groups exploit the weak and don't pay their fair share of taxes. Conservatives distrust "Big Government," which takes money from hardworking Americans and gives it to lazy people and unauthorized immigrants.

In a strictly political context, on the left, fairness often means equality. On the right, fairness often means proportionality (i.e. people should be rewarded in proportion to what they contribute).

In normal life, fairness usually means proportionality. When people work together on a task, studies show they are more concerned about equality of outcome if they know that everyone gave equal effort. If they think that some people contributed more than others, people generally believe those big contributors deserve more gains. This comes from people's strong desire to protect their communities from cheaters, slackers, and free riders who would cause others to stop cooperating and cause society to unravel if not stopped from doing so. Hence much of the right's obsession with the social safety net, which might allow some to mooch off government's assistance.

### The Loyalty-Betrayal Foundation

Humans are not the only species that wages war or kills its own kind. Chimpanzees guard their territory, raid the territory of rivals, and kill off males of the neighboring group to steal their territory or their females. Warfare among humans has been a constant feature of human life for millions of years—long before agriculture and private property. Many psychological systems contribute to effective tribalism and success when groups are locked in inter-group competition. The loyalty foundation is just one part of our innate ability to form cohesive coalitions. The trigger for the loyalty foundation is anything that tells you who is a team player and who is a traitor.

### The Authority-Subversion Foundation

The desire to show respect is common across animal species. There are dominance hierarchies in chickens, dogs, chimpanzees, and many others. Physical displays are made by low-ranking individuals to higher ones. They usually involve similar gestures, like appearing small and non-threatening.

Human cultures vary enormously by how much respect they demand for parents, teachers, and others in positions of authority. The urge to respect hierarchical relationships is so deep, many languages encode it directly into their grammar structures. For example, French speakers must choose between the respectful form (*vous*) and or the familiar form (*tu*) when speaking to another person. English speakers don't use this construction, but you may have felt a pang of awkwardness when an older person you have long revered asks you to call her by her first name.

The authority foundation helps individuals forge beneficial relationships within their tribe's hierarchy. Individuals best able to rise in status while cultivating the protection of

superiors and allegiance of subordinates survive best, and are best able to pass on their genes.

Like chimps, humans track and remember who is above whom in the tribe's hierarchy. When people within a hierarchy do things that upend that order, we notice it instantly. Since authority is (in part) about protecting order and fending off chaos, everyone has a stake in upholding that order. That means holding people accountable for fulfilling the obligations of their place in the hierarchy.

Authority is not the same as power or control. Even among chimps, the authority is usually an alpha male who performs some socially useful functions, like resolving disputes and suppression much of the violence that erupts without a clear leader.

Among humans, authority figures take on responsibility for maintaining order and justice. (They also may exploit that power for their own benefit).

The current triggers of the authority foundation include anything that can be seen as an act of obedience or disobedience, respect or disrespect, submission or rebelling, when the authorities are perceived to be legitimate. This includes metaphorical authorities like traditions, institutions, or values that people believe provide stability.

### The Sanctity-Degradation Foundation

Most animals are born knowing what they should eat. For example, the koala bear has an instinct to only eat eucalyptus leaves. Humans, though, don't have this instinct because we are omnivores—we eat a very wide variety of things. While this gives us flexibility, it can cause us to eat toxic food, diseased food, or food with parasites.

For this reason, omnivores go through life loving new foods and being afraid of new foods. Individuals will fall somewhere on a spectrum between loving novelty (which benefits one's survival) and fearing novelty (which also benefits survival).

Humanity's tolerance (or intolerance, depending on the person) for novelty carries over into political views. Liberals, for example, score higher on psychology measures of their openness to experience new food, music, and ideas. Conservatives score higher on the opposite, which is why they care more about guarding borders and traditions.

It makes evolutionary sense why openness to new foods would carry over into views on how to order the social world around us. When early humans began to live in larger groups, they greatly increased their risk of infection from each other. They developed what Mark Schaller has called the "behavioral immune system"—a set of predispositions that are triggered by signs of infection, disease and parasites, which make us want to avoid touching the person or thing carrying the disease. Key triggers are smells, sights, or other sensory patterns that predict the presence of dangerous pathogens (such as human corpses, excrement, vultures, and people with sores). Current triggers of the Sanctity foundation are extraordinarily variable. For example, immigrants (who might bring new plagues, epidemics, and diseases) are more favorable when disease is less common.

The Sanctity foundation explains why people treat objects (e.g. flags, crosses), places (e.g. Mecca, a battlefield), people (e.g. saints, heroes), and principles (e.g. liberty, fraternity, equality) as though they had infinite value. When someone desecrates one of the sacred pillars supporting the community, the reaction is swift, collective, and punitive. Philosopher Leon Kass believes that the Sanctity foundation can help humans



avoid going too far when transgressing long-held taboos, even when we cannot rationally justify the old taboos by pointing to victims. As he once wrote, “Shallow are the souls that have forgotten how to shudder.”

Today, the Sanctification foundation is more apparent among the religious right. But the spiritual left also has its New Age grocery stores stocked with products that cleanse your “toxins.” The Left also has the environmentalists, who revile industrialism, capitalism, and cars for both their real pollution and their symbolic degradation of nature.

## Chapter 8: The Conservative Advantage

Haidt, along with Jesse Graham and Brian Nosek, created the first Moral Foundations Questionnaire to understand how American liberals and conservatives’ moral foundations differed. They asked people 15 questions—three of which related to each of the five foundations. For example, Care questions would be along the lines of, “Whether or not someone was cruel”; Authority questions would ask “Whether or not someone showed a lack of respect for authority.” Every participant, regardless of their political views, said that concerns about compassion, cruelty, fairness, and injustice are relevant to their judgments about right and wrong.

But liberals say these issues are more relevant to their judgments about right and wrong. Conservatives care much more about the other foundations—Loyalty, Authority, and Sanctity—than do liberals (who largely reject them as relevant to morality). Conservatives don’t dismiss Care and Fairness, they just tend to think these values are roughly equally important as Loyalty, Authority, and Sanctity. Liberals are almost exclusively concerned with Care and Fairness.

Haidt and his colleagues’ findings were replicated even when the researchers asked people about the traits of their preferred dog breeds. Liberals were more interested in Caring breeds and breeds that treated their owners as equals. Conservatives were more interested in breeds with traits that balances all five moral foundations.

These findings were replicated in people’s church affiliations. The researchers applied a word processing program to Unitarian (i.e. liberal) church sermons and Southern Baptist (i.e. conservative) church sermons to detect words that relate to each moral foundation. For example, peace, care, and compassion are related to Care; obey, duty, and honor are related to Authority. They found that the Unitarian preachers used Care and Fairness words more than Baptist preachers, while Baptist preachers used words relating to Loyalty, Authority, and Sanctity more.

These findings were replicated in brain scans. Liberals and conservatives were asked to read certain sentences that either supported a moral foundation or opposed it. Liberals showed more brain activity related to shock and surprise when they read sentences that rejected the Care and Fairness foundations, or that supported the Loyalty, Authority, and Sanctity concerns.

## The Liberty-Oppression Foundation

Haidt and colleagues added a 6<sup>th</sup> foundation to make sense of data on people’s reaction to Fairness questions. Liberals seemed to think of fairness as equality. Conservatives seemed to think of fairness as proportionality. Why?

Haidt's explanation is that humans evolved in nomadic hunter-gatherer tribes that were highly egalitarian and had no hierarchies. The norms of the group actively encouraged sharing. Hierarchy only became widespread when groups began to start farming and domesticating animals. This economic change created "private property" for the first time. Human groups also got much larger. Equality ended.

Christopher Boehm has researched hierarchies in chimps and humans and found that humans are innately hierarchical, but not as much as chimps. Chimp authority figures, for example, provide few useful social functions beyond mitigating violence among the troop. Rather, chimp authorities usually just act like bullies, taking food and females from weaker chimps. Sometimes chimps band together to assassinate the alpha male if he goes too far. Chimp leaders, therefore, have to balance their power and know how far they can go.

Boehm concludes that humans became less hierarchical in the last one million years once invented weapons for hunting and language. This gave the non-alpha males the ability to gossip about the alpha male (i.e. rally against him) and kill him (with their new weapons) even if they were weaker than him. Humans could unite to shame, ostracize, or kill the bully alpha male, whose behavior threatened or annoyed the rest of the group. Hierarchies collapsed and human tribes became much more egalitarian.

In this new world, people were quick to develop richer moral matrices of norms, informal sanctions, and occasionally violent punishments. Those who could navigate these new cultures advanced socially; those who couldn't were punished or exiled. Such is how humanity became more cooperative and sociable—what Boehm calls "self-domestication."

All this history gave humans an inborn sensitivity to individuals who, if given the Chance, would dominate, bully, and constraint others. One person's signs of domination would trigger righteous anger in others. If groups didn't react this way, they would be exploited by one domineering alpha figure and would be less cooperative, and less successful, as a tribal unit. Hence why people have a moral inclination for "freedom."

The moral foundation for Liberty has bifurcated into two expressions. Liberals sacralize *equality*, which is pursued by fighting for equal civil and human rights and for equal outcomes. Conservatives sacralize negative liberty, freedom from the nanny state and high taxes, from oppressive regulations, and from the United Nations and other sovereignty-reducing international treaties. Notice again that conservatives interpret their moral matrix in terms of how it protects their group, whereas liberals interpret the foundation more universally.

## Part III: Morality Binds and Blinds

*"We are 90% Chimp and 10% Bee."*

### Chapter 9: Why Are We So Groupish?

People are "*groupish*." We love teams, club, leagues, and fraternities. People are great at promoting our group's interests when competing against other groups. People,

though, are also *selfish*, which means we are good at promoting our own individual interest when that conflicts with the interests of our peers. Haidt believes human nature is mostly selfish, but with a groupish overlay.

For a long time, scientists didn't believe groupish-ness was possible among groups that weren't genetically related. For example, the bravest army wins, but the soldiers most likely to survive the battle and go home to have children are the cowards that hung back and avoided the most dangerous fighting. In the words of evolutionary biology, a gene for suicidal self-sacrifice would be favored by the group-level natural selection but be strongly opposed by individual-level selection.

But it is possible to develop groupish instincts through a multi-step evolutionary process. Step 1: Individuals with social instincts to stay close to other humans would become more likely to avoid being eaten by predators. Step 2: People who helped others were more likely to get help when they needed it most. Step 3: Language gave people the power to share information about others (i.e. gossip) which created reputations. Step 4: Groups treated duties to the group and principles as sacred, enforcing observance of these duties through gossip and social sanction.

Once these four steps take hold, then the free riding soldier could not go to battle and hide in the back. Stories of his cowardice would make it back to the tribe, and he would be beaten up by his fellow survivors, become repellent toward women at home, and be generally shunned.

Finding a way to suppress free riders has been the central theme of evolution. In the beginning of life one billion years ago, there were only one-celled bacteria. When two bacteria combined, they became a cooperative unit that could divide labor, be more efficient, and outcompete other bacteria cells. Then, bacteria grouped together into multi-celled bacteria; division of labor increased even more, and multi-celled organisms easily outcompeted their one-celled competitors. Multi-celled organisms developed complex, inter-dependent systems, creating plants, animals and fungi. When humans formed cohesive groups, it was another example of individual entities coming together to form a super-entity. Haidt explains: "Whenever a way is found to suppress free riding so that individual units can cooperate, work as a team, and divide labor... selection at the higher level becomes more powerful, and that... favors the most cohesive superorganism."

Many animals are social, but only a few are *ultrasocial*. This select group includes humans, ants, bees, among a few others. These ultrasocial animals tend to have similar characteristics that lead to their ultrasociality. They often share a nest, which must be protected from predators, parasites, or competitors; the need to feed offspring over an extended period; and they often face lots of conflict with other groups of the same species. For example, wasps have nests (that is, holes in trees) and have to compete with other wasp colonies nearby.

Humans also exemplify these traits. As hunter-gatherers, we were once territorial creatures with defensible nests, like caves. We gave birth to needy offspring that required enormous amounts of care. And when the tribe was under threat by another group of humans, there was an intense need to bind together. Agriculture made our homes more permanent and more valuable, which only ratcheted up the need to band together to defend the farmland and other tribal resources. Soon, competitive pressures led to city-states, walled cities and armies, and eventually empires.

Among ultrasocial creatures, only humans have gone this far. For example, even our closest cousin—the chimp—will never cooperate with another chimp to carry a log, despite being the second-smartest species on the planet (they are able to make tools, learn sign language, predict the intentions of other chimps, and deceive others to get what they want).

Michael Tomasello has done experiments to tease out what sets humans apart from chimps. He compared chimps to two-year-olds, asking them to perform simple tasks like use a stick to pull in a treat that was out of reach. At these tasks, chimps were as good as any two-year-old human. The difference between the chimps and humans, though, was that humans have what Tomasello calls “shared intentionality.” Humans can imagine what the other person is thinking and use that awareness to coordinate our behavior. When everyone in a group begins to share a common understanding of how things are supposed to be done, they can cooperate on incredible complex tasks. Once we gain this ability, humans can learn and conform to social norms, feel and share group-related emotions, and ultimately create social institutions like religion. This unlocked the next level of evolution—competition between groups of ultrasocial, cooperating animals.

Once humans started competing at the group level, we imposed a new evolutionary pressure on ourselves which started to change us. We developed an obsession with symbolic markers to show our group memberships. For example, today we can see people tattoo and pierce their face to show which Amazonian tribe they belong to; we see Jews circumcise their males; we see tattoos and facial piercings by punks in the UK. These symbols tell us who we can trust and cooperate with easily, because just by looking at them we know they share our values and norms. We then developed more psychological traits, like an expectation that life is structured by moral norms. We developed new emotions to internalize those norms, like shame, guilt, and righteous anger. As humans continued to select friends and partners based on their ability to live within a tribe’s moral matrix, these characteristics became stronger and stronger over time. Like other domesticated animals, this process of “self-domestication” led us to have smaller teeth, smaller bodies, reduced aggression, and greater playfulness (even into adulthood).

Genetic evolution as outlined above can happen relatively quickly when the pressure is just right. Dmitri Belyaev showed that with individual selection for the most docile and social fox, foxes can be completely domesticated within about 9 generations. William Muir showed that using group selection to breed hens (i.e. breed the *group* of hens that collectively produced the most eggs, as opposed to the *individual* hen that produced the most eggs) required 6 generations to reach a 260% increase in egg production.

Humans might have taken more than 6 to 9 generations to respond to their new group level selection, but they sure changed a lot genetically in a short period. The human genome project, for instance, shows that genetic evolution greatly accelerated during the last 50,000 years. High rates of climatic fluctuation (from hot to cold, dry to wet, and back again) correlates with the fastest period of genetic change in humans. Between 70,000 and 20,000 years ago, there were several periods where almost all humans were killed. During these phases, those groups facing imminent death but who could cooperate in sophisticated ways to monopolize, hide, and share scarce food

supplies were the ones that survived. Their offspring would become our modern day, ultrasocial, ultra-cooperative selves.

## Chapter 10: The Hive Switch

Haidt believes that humans can be “hive” creatures under some conditions. That is, humans transcend their self-interest and lose themselves temporarily and ecstatically in the larger group—if their hive nature is activated. Haidt offers three common examples of hive emotionality that you might identify with (or know someone who identifies with): 1) Awe in nature; (2) Hallucinogens like mushrooms, peyote, ayahuasca, and LSD, all of which can shut down people’s experience of themselves and give them a “religious” or “transformative” experience; and (3) Raves, where electronic music with hypnotic melodies, trance-inspiring light shows, and heavy basses create a communal feeling among dancers.

The neurochemical that binds people together, oxytocin, creates a powerful motivation to touch and care for one’s children. In species in which males stick by their mates or protect their offspring, it is because their brains are more responsive to oxytocin. If you squirt oxytocin into someone’s nose, he will be more trusting in a game that involves temporarily transferring money to an anonymous partner. Oxytocin levels rise in people who are treated with trust by others.

Several studies have shown, though, that oxytocin only binds people of the same group, not humanity in general. A study on Dutch men, for example, found that when the men played a variety of economic games while sitting alone in cubicles, those who had oxytocin sprayed in their noses made less selfish decisions and more beneficial decisions for their group. They showed no concern at all, however, for improving the outcomes of men in other groups. In a follow-up study, oxytocin made participants like those with Dutch-sounding names more and value saving Dutch lives more. This line of research has never shown oxytocin increases out-group hate—it simply makes people love the in-group more than they did before.

Mirror neurons also help bind people together. Mirror neurons light up when we watch someone else do something. Watching a hand pick up a cup, for example, triggers a mirror neuron for eating. Seeing someone else smile triggers a mirror neuron to smile. Our mirror neurons light up more when we observe people we like. In one experiment Tania Singer had participants play an economic game with two strangers, one who played nice and the other who played selfishly. In Phase two of the study, participants watched the nice player and the selfish player receive electrical shocks. Subjects’ brains showed evidence of pain when the nice player was shocked, but no evidence of pain when the selfish player got shocked (some even showed pleasure). As Haidt writes, “We are more likely... to empathize with others when they have conformed to our moral matrix than when they have violated it.”

All of this kicks out some basic advice for leaders who want to create more hive-like super-organizations. First, they can construct a moral matrix that emphasizes certain oft-forgotten values. The organization’s culture should legitimize the authority of the leader in order to appeal to the Authority moral foundation. The culture should make sure subordinates don’t feel oppressed by appealing the Liberty foundation. And it should appeal to the Loyalty foundation.

Second, leaders should emphasize similarity—not diversity. Don't call attention to racial and ethnic differences. A great deal of research shows that people are warmer to and more trusting of people who look like them, dress like them, talk like them, or even just share their first name or birthday. Drawing attention to these similarities will foster cooperation more than appealing to differences.

Third, leaders should exploit synchrony. If you ask people to sing together, march in step together, or just snap out some beats together, this will make them more trusting of each other and more likely to help each other out.

Finally, leaders should create healthy competition among teams. Studies show that intergroup competition increases love for the in-group far more than it increases dislike of the out-group. So group vs. group competition is a net positive for the whole.

However, pitting individuals against individuals in a competition for scarce resources (like bonuses) will destroy hive-ishness, trust, and morale.

## Chapter 11: Religion is a Team Sport

One theory of religion is that it is a cultural innovation that makes groups more cohesive and cooperative. Groups with “less effective” religions don't necessarily get wiped out. But they often just adopt the more effective variation. In this way, religions spread like technology.

This theory comes from the observation that religions seemingly “evolve” over time. For instance, hunter-gatherers often had religions where the gods were capricious and malevolent. The gods sometimes punished bad heavier, but would also cause the righteous to suffer, too. Agricultural societies, on the other hand, tended to have far more moralistic gods. Their religions were focused on condemning murder, adultery, lying, and breaking promises. Their gods administered collective punishment, such as disease, pestilence, or drought, on a whole village for the adultery of two people. Beliefs like this would make villagers much more vigilant and gossipy about any of their neighbor's slightest deviation from the community's moral matrix.

Evidence shows that moralistic religions, such as those that originated during humanity's agricultural revolution, help groups cohere, solve free rider problems, and win out against other groups with whom they compete. Richard Sosis examined the history of 200 communes in the US during the 1800s. Communes are usually colonies in the wilderness founded by a committed group of believers who reject society's moral matrix and go off to form their own society. He compared communes that were religious to ones that were secular. Just 6% of secular communes lasted longer than 20 years. Meanwhile, 29% of religious communes lasted longer than 20 years. Sosis quantified everything he could about those communes to find out why the religious ones were doing better. He found that the religious communes were more likely to expect costly sacrifices from members, like giving up alcohol and tobacco, fasting for days at a time, conforming to a communal dress code, cutting ties with outsiders. The correlation was perfectly linear—the more sacrifice a religious commune demanded, the longer it lasted. Interestingly, the more sacrifices demanded by the secular communes did not increase their odds of survival. Sosis explains that this is because in secular communes, members would be more willing to question why the group demanded the sacrifice than the religious commune members would. They might even refuse to adopt a sacrifice

that “didn’t make sense.” The religious communes were better at binding people together because they used the pretense of sacredness to blind the members to the arbitrariness of the “sacred” practice.

In another example of religion fostering cooperative societies through strict moral rules, John Sloan Wilson has showed how John Calvin developed a strict, demanding form of Christianity that suppressed free riding and facilitated trust and commerce in Geneva in the 1500s. He’s also shown how medieval Judaism created a “cultural fortress” that kept outsiders out and insiders in.” he’s shown how Balinese rice farmers solved a very difficult free rider problem related to scarce water supplies by placing small temples at the fork in the irrigation system. The god of each temple united all the rice farmers together in a community worshipping that god, helping them solve their disputes more amicably and minimizing cheating and deception. As Haidt explains, “Gods are tools that let people bind themselves together in a community by circling around them.” People belonging to religious societies are more likely to survive and reproduce, meaning the next generation will be more religious than the last. Such is how religion became an indelible feature of humanity. Gods and religions are group-level adaptations for producing cohesiveness and trust.

We can see this in laboratory studies. Studies show that people who find out you are religious will trust you more with money. And those who are religious will share back more money to a person who trusted them than non-religious people will (regardless of whether the religious person knew they were returning money to a religious or non-religious other).

We can also see this in history. Jews and Muslims long excelled in long-distance trade, in part because their religions helped to create trustworthy relationships and enforceable contracts. Even today, markets that require a high deal of trust (like diamond markets) are dominated by ethnic groups (like ultra-Orthodox Jews) who have lower transaction and monitoring costs than their secular competitors.

Sociologists Robert Putnam and David Campbell summarize religion’s effect on society like this: “By many different measures, religiously observant Americans are better neighbors and better citizens than secular Americans—they are more generous with their time and money, especially in helping the needy, and they are more active in community life.” When Putnam and Campbell peered deeper into what it was about religion that made people act better, they found it was not belief in hell, or daily prayer, or belonging to a certain Christian sect. it was the fact that religion bound people into relationships with their fellow worshippers. They concluded, “It is religious belongingness that matters for neighborliness, not religious believing.”

## Chapter 12: Can’t We All Disagree More Constructively?

Political theorists long thought that people chose political ideologies that benefit their self-interest. For example, the rich would want to preserve the existing order and the peasants would want to change everything. But in modern times, political scientists have found that self-interest does a remarkably poor job of predicting political attitudes. The next theory political scientists cooked up was that people soak up political ideologies from their parents or the TV programs they watched growing up.

But studies of twins found that political ideologies can be largely linked to genetics. Identical twins, for instance, share 100% of their DNA, while fraternal twins share 50% of their DNA. This means that twins are far more similar in terms of IQ, mental illness, basic personality traits, tastes in music and food, their propensity to get divorced as adults, their likelihood of dying in a car crash, their degree of religiousness, than are fraternal twins. And when it comes to political views, they are also more similar to each other than fraternal twins are to each other. This suggests that DNA shapes political views.

This finding is strong enough to explain between one-third to one-half of the variation among people's political views. Being raised by conservative or liberal parents explains far less.

Genes influence your political ideology in two steps—(1) brain biochemistry and (2) personality development. First, the genetic difference between political liberals and conservatives seems rooted in their neurotransmitters—specifically, glutamate, serotonin, and dopamine. Glutamate and serotonin neurotransmitter are involved in the brain's response to threat and fear. This aligns neatly with other findings that conservatives react more strongly to signs of danger, like germs or loud noises. Dopamine, meanwhile, is related to people's sensation-seeking and openness to new experiences, which are behavioral traits correlating strongly with politically liberal attitudes.

Second, personality has levels. Per Dan McAdams, there are three levels. Our lowest level includes our "dispositional traits" that generate our general characteristics—are we thrill-seeking or nervous about threats, extraverted or introverted, conscientious or easy-going? These are determined by the neurotransmitters described above. Then, there's the second level, our "characteristic adaptations," which we learn along the way to help us fit into certain environments. If our genes steer us toward different experiences as we grow up, these experiences then reinforce and tailor our original genetic dispositions. For example, the cautious brother remains close to his hometown and gets involved in church, while his sister moves to New York City and becomes an advocate for unauthorized immigrants. Both find themselves in different social environments, and mature in ways that help them navigate those environments. Our third level is our "life narrative," which is the story we tell ourselves about ourselves to help make sense of our life trajectory. Narratives don't have to be literally true—they are selective and simplified, often for the sake of creating an idealized vision of the future. They are saturated with morality. Although these narratives can be after-the-fact fabrications, they still influence people's behavior, relationships, and mental health. Life narratives fuse our identity to the moral matrix of our society, which we gravitate toward because of our low-level influences (genes) and second-level influences (acclimation to our social niche).

It's worth elaborating on how individual narratives blend into larger group narratives, which we'll call "Grand Narratives." Each society (i.e. group of people regularly interacting with each other), puts at the core of its Grand narrative a sacred something or other. Then society constructs a story with a beginning, a middle (where the "threat" arises) and an end (in which a resolution is achieved). The narratives orient listeners morally, drawing their attention to a set of virtues and vices, good and evil forces, to



impart lessons about what must be done now to protect, recover, or attain the sacred core of their vision.

Christian Smith has cataloged many Grand Narratives of moral tribes living in modern America. For example, he offers the following Liberal Progress Narrative:

“Once upon a time, the vast majority of human persons suffered in societies and social institutions that were unjust, unhealthy, repressive, and oppressive. These traditional societies were reprehensible because of their deep-rooted inequality, exploitation, and irrational traditionalism... But the noble human aspiration for autonomy, equality, and prosperity struggled mightily against the forces of misery and oppression, and eventually succeeded in establishing modern, liberal, democratic, capitalist, welfare societies. While modern social conditions hold the potential to maximize individual freedom and pleasure of all, there is much work to be done to dismantle the powerful vestiges of inequality, exploitation, and repression. This struggle for a good society in which individuals are equal and free to pursue their self-defined happiness is the one mission truly worth dedicating one’s life to achieving.”

Then there’s the Reagan Republican Narrative of the 1980 election:

“Once upon a time, America was a shining beacon. Then liberals came along and erected an enormous federal bureaucracy that handcuffed the invisible hand of the free market. They subverted our traditional American values and opposed God and faith at every step of the way... Instead of requiring that people work for a living, they siphoned money from hardworking Americans and gave it to Cadillac-driving drug addicts and welfare queens. Instead of punishing criminals, they tried to “understand” them. Instead of worrying about victims of crime, they worried about the rights of criminals... Instead of adhering to traditional American values of family, fidelity, and personal responsibility, they preached promiscuity, premarital sex, and the gay lifestyle... and they encouraged a feminist agenda that undermined traditional family roles... Instead of projecting strength to those who would do evil around the world, they cut military budgets, disrespected our soldiers in uniform, burned our flag, and chose negotiation and multilateralism... Then Americans decided to take their country back from those who sought to undermine it.”

If the stories are so different, can each side even understand the other? Haidt, along with Jesse Graham and Brian Nosek, tested how well liberals and conservatives could understand each other. They asked over 2,000 Americans to fill out the Moral Foundations Questionnaire. They specifically asked one-third of participants to answer

the questionnaire as they normally would; they asked one-third to answer like a “typical conservative” would; and they asked one-third to answer as a “typical liberal” would. This allowed Haidt, Graham, and Nosek to test how accurate participants were in guessing their political opponents’ views by comparing people’s expectations about the “typical” partisan view to the actual responses from conservative and liberal partisans. They found that moderates and conservatives were more accurate in their predictions of what the typical liberal or typical conservative believed. Liberals were the least accurate—especially those who described themselves as “very liberal.” The biggest errors came when liberals tried to guess what the typical conservative thinks about care and Fairness questions. They assumed conservatives would disagree with the following statements:

- “One of the worst things a person could do is hurt a defenseless animal” (Care foundation)
- “Justice is the most important requirement for a society” (Fairness foundation)

Haidt believes liberals in American politics have another blind spot: They miss the importance of social capital and moral capital. Conservatives believe that people are inherently imperfect and are prone to act badly when all constraints and accountability are removed. But through institutions, we can account for this weakness through moral capital. Moral capital (i.e. the general commitment by everyone to do the right thing and not be selfish when no one’s looking) is generated among communities with strong social capital (i.e. the number and strength of relationships between individuals, and people’s commitment to reciprocity and trustworthiness to those they have a relationship with). Social capital could take the form of formal membership to institutions, like university or church. Or it could be the informal ties of shared culture, traditions, and identities that make people similar to each other. Social capital encourages people to adopt similar norms of good behavior and reinforces those norms with formal laws and religions as well as informal customs and traditions.

Liberals build their moral matrix on Care, Fairness, and Liberty (where Liberty refers to equality of outcomes, which maintains parity, which prevents domination). These three-prong foundation helps them see the value of restraining super-organisms like corporations that threaten the weak with their awesome accumulation of power. This helps liberals address large-scale regulatory problems, like lead gasoline (which Democrats phased out through regulation in the 1970s), and whose effects require collective effort and fall disproportionately on the vulnerable.

Libertarians build their moral matrix almost exclusively on Liberty (where Liberty here means freedom from coercion). This helps libertarians see the value of unrestrained free markets, which provide the enormous plenty that freed humanity from subsistence-level poverty that trapped us and limited our potential for almost all of human history.

Conservatives build their moral matrix on all 6 foundations, which means they are the only major political ideology to emphasize Loyalty, Authority, and Sanctity. These extra foundations help conservatives appreciate the value of social capital and moral capital for uniting us against collective social problems. Robert Putnam, for instance, has found that American communities with high levels of immigration and ethnic

diversity seemed to cause a reduction in social capital. Was it because people are racist, such that they withdrew from their diversifying community? No, says Putnam. Instead, he found evidence that trust between groups and trust within groups deteriorated. That is to say, instead of binding groups to their own racial compatriots and alienating them from their new, diverse neighbors, what actually happened was that everyone distanced themselves from everyone, even their own ilk. Individuals just retreated from social life altogether. The threat diversity brings to social capital is something overlooked by liberals and libertarians, who failed to appreciate the ties that bind people together, and the good that those ties between people (i.e. social capital) then go on to create for society.

# The Ruin of Kasch by Roberto Calasso

Roberto Calasso, *The Ruin of Kasch*. Translated by William Weaver and Stephen Sartarelli. Belknap Press, Cambridge University, 1994. (Italian original 1983) Reviewed by John Bedell

## The Revolution

A character in one of Saul Bellow's novels comments that he hates facile "explanations" of the Holocaust; such an event, he says, could be explained "only by explaining everything at once." In *The Ruin of Kasch* Roberto Calasso has a go at doing exactly that.

*The Ruin of Kasch* (1983, English translation 1994) is a maddening, intentionally difficult book about what the human world is like, what it used to be like, and how it changed. Calasso is not interested in a statistical, factual understanding of history, nor in any kind of linear story. Instead he gives us a strange mixture of quotation, gnomic utterance, striking juxtapositions, narratives of obscure incidents, brief biographies of the moderately famous, and cutting analysis of various theories that purport to explain something. His attention jumps around in time and space, with no transitions or explanations. His examples are sometimes so obscure that even experts in whatever period he is writing about may not recognize them. Some of the book makes no sense. Yet by the end I at least was left feeling that I understood history in a way that I had not before.

Calasso begins with the era of the French Revolution. But this, perhaps the only conventional choice in the whole book, is immediately twisted by his decision to focus his narrative on the character of Talleyrand. Talleyrand, a son of one of France's oldest families, was a survivor, a relic of the old regime who prospered through all the transitions and upheavals of the epoch and emerged at the end as a kingmaker and one of the most powerful and hated men in Europe. Why Talleyrand? Because, I think, of his cold distance from the events he helped to shape. He helped create the new rituals of the Democratic age – tricolor flags, national anthems, independence days – but they made him queasy. Contemplating the future from the first celebration of Bastille Day, he wrote, "I see streams of blood." (84)

In the short term the Revolution was cut short by reaction, dictatorship, empire, and wars that dwarfed any the continent had seen in centuries. Instead of the Brotherhood of Man, Europe got the great divide between Left and Right that is still with us. After Waterloo the forces of reaction managed to bottle up the genies of revolution and nationalism for a time, but it could not last.

## The Tale Told in the Desert

From the Age of Revolutions Calasso jumps with no transition to the legendary past, retelling an old story called "The Ruin of Kasch". The tale was recorded by the German anthropologist Leo Frobenius in the early 1900s, told at the edge of the Sudanese desert by an old camel driver who waited to tell his own tales until all the other, lesser speakers had exhausted themselves.

You know nothing of stories, he said to them, before embarking on the tale of Naphta, a grand ancient kingdom ruled by the guidance of priests who studied the stars. They decided when the king would be sacrificed and a new king crowned. Each king therefore ruled only on their sufferance, awaiting the day when the stars would decree his death. One privilege the kings had was that of choosing their companions, who would accompany them in life and in death.

A certain king named Akaf grows sad and cannot stop brooding on his upcoming death. He hears about a great storyteller known as Far-li-mas who came from Kasch in the far east, perhaps Arabia or India. Akaf summons Far-li-mas, who tells such wonderful stories that the dawn comes before anyone realizes that the night is passing, and the king forgets his sorrow. The king's sister Sali also hears the stories, and she and Far-li-mas fall in love. Unwilling to leave their fate up to the stargazing priests, Sali and Far-li-mas challenge the old way. Sali says,

Great are the works of God, but the greatest is not his writing in the sky. It is life on earth. (120)

Sali lures the priests to the court he hear Far-li-mas, and he tells his stories like hashish, so that all the listeners fall asleep, and the priests cannot watch the stars. After some nights the priests realize that they are losing track of the stars and tell the king that Far-li-mas has destroyed order and must be killed. The king summons all the people to the great square of the city so that God may decide the matter. Again Far-li-mas tells his stories, and in the morning the priests are all dead. The old way of sacrifice is abandoned. King Akaf lives until he dies a natural death, and Far-li-mas reigns after him with Sali as his queen.

But upon the death of Far-li-mas, the neighboring kingdoms abandon the oaths of friendship they had sworn to the great Akaf and make war on Naphta. The kingdom is destroyed, the cities abandoned, and the desert covered it once broad fields of grain.

Calasso tells us,

This is a story about the passage from one world to another, from one order to another—and about the ruin of both. It is the story of the precariousness of order: of the old order and the new. The story of their perpetual ruin. (139)

Although Calasso is not impressed by modernity, he is equally unenthralled by the ancient world. No society has really found a solution to the problems of being alive. What he seems to hate about modernity is that we keep proclaiming that we have, in fact, done so, while the ancients at least realized that this was beyond our powers.

## Sacrifice

Blood sacrifice is one of Calasso's constant themes, and he often makes it a sign of the difference between us and our ancestors: they carried out sacrificial rites, and we do not. This might seem odd, since blood sacrifice disappeared from Europe more than a millennium before the modern age. When I first read *The Ruin of Kasch* I did not understand where he was going, but after reading his other books and coming back to this one I have some theories. For one, Calasso makes the rite of sacrifice stand in for all the ways we destroy in order to live. He

sometimes describes slaughterhouses as sites of sacrifice on an industrial scale, and he writes that with World War I human sacrifice returned to Europe on a grand scale.

More subtly, Calasso seems to regard sacrifice as a way of defining human consciousness, not neurologically, but in terms of experience. The sacrificer beholds the other, the animal to be killed, and recognizes it as something separate from himself; and yet they are also connected, since the victim is actually standing in for the sacrificer, one life for another; he is thus aware that one thing can stand for another, as a word for a thought; he feels guilt over the death he must cause, and yet feels also that it must be done; he thinks of other sacrifices, those he remembers and those he knows of from stories, including stories of how the gods first established sacrifice at the dawn of history; he feels himself part of this long chain of killings; yet he is alone with the knife in his hand, confronting a deed he must do. When the animal is killed there is at first a *katharsis*, a release of tension, but after that comes more guilt that must be atoned for with yet more sacrifices, creating an unending chain of killing and guilt. Sacrifice, Calasso writes, does not expiate guilt; it is guilt.

Consciousness is built up out of these constant recursions, the thoughts and feelings that loop back on themselves over and over, from the self to the other and then back to the self, bringing in memory and myth and connection and separation. Sacrifice is an act of separation, a killing, but it is also a connection that ties the sacrificer to the divine. In terms of history, we might say that the ancient attitude made consciousness a holy thing, consecrated to the gods through sacrifice, through all the loops of thought that run through divine law and divine story. In this sense Calasso's model is related to that of Julian Jaynes in *The Origin of Consciousness in the Breakdown of the Bicameral Mind*: the Gods used to walk beside us, but we left them behind. Calasso differs from Jaynes in refusing to assign this change to any particular point in time; to Calasso it is still taking place, in all of us.

The wise man, he suggests, can accept neither the religious postulates of the ancients nor life without them. On the one hand there are the fantastic rituals of the Indian Vedas, too complex to ever be enacted in every detail, yet all said to be essential for existence to continue. On the other there is only Bentham's utilitarianism, represented to Calasso by his dried-up mummy, still kept in London. One cannot be believed, the other is inadequate to our needs. "We are in the middle, wavering."(173)

## Goethe Beholds the Paintings

Here is one of the stories Calasso tells us, without much in the way of preamble or explanation, as if saying, "make of this what you will." It concerns the wedding of the Austrian princess Marie Antoinette to the future king of France, but not as a political or diplomatic event. Calasso instead describes the pavilion built in the middle of the Rhine at the exact boundary of French and Austrian territory, where Marie would be handed over from one nation to the other. There she was stripped completely naked and then dressed anew in clothes made entirely in France before she passed into her new home. As it happens the poet Goethe visited the pavilion before the princess passed through, and he observed the elaborate paintings that covered every wall. To his horror, he realized that they depicted the story of Jason and Medea. This is monstrous, he exclaimed to his companions; how can this wedding pavilion be painted with scenes of history's most disastrous marriage? There is no need to worry, said his companions, nobody pays any attention to the subject matter of paintings. Only the style matters.

## The Young Hegelians

Already in the heady days of the French Revolution we meet one of the key figures of modern history, the metaphysicians of Terror, the men who believed that utopia was within reach if we could only remove the enemies blocking the way. Calasso cites several of these men, most of them utterly obscure, but perhaps more frightening because they were otherwise such ordinary people. Like a certain Monsieur Baudot, who wrote,

The egoists, the thoughtless, the enemies of liberty, the enemies of all nature, must not be counted among her children. . . . Let us destroy them completely.  
(167)

One of these writers called for the extermination of a third of the population. People at the time saw this as an old enemy, religious fanaticism, with the people in place of God, and the nation standing in for the church. But the language was new, and the belief that utopia could be built by human hands, without divine aid.

Two generations later a new wave of utopian thinkers emerged from among the students and followers of Hegel. Marx is the most famous, but there were many others, young men who thought that Hegel had taken philosophy as far as it could go as thought: the only way to move forward was through action. Philosophy must shift, one wrote, "to the absolutely practical terrain of the will." Through the will, philosophy will achieve, as Bakunin put it, "a complete reconciliation with reality in every area of life." Calasso calls this the "fatal shard of the Hegelian legacy." (256) The world – that is, society – must be made to conform to philosophy in every particular.

If Calasso were in favor of cancelling things, which he is not, he might have inserted here: if we are going to silence anyone, it should not be the trolls or the cruel jokesters but the suave philosophers who want to twist the world until it fits their visions of perfection.

## The Nihilists

At around the same time that the young Hegelians were turning philosophy from thought into "praxis," the other great strain of modern thought received new impetus: nihilism. Nihilism was given its most perfect form early on by Stirner, a contemporary of Marx, who wrote that all metaphysics was "mad raving," and that nothing really existed but the brute facts of animal life. Everything else was a mere "ghost" – "spook," his most recent English translator renders this – a story told to frighten us and keep us from knowing the truth. The stage is thus set for two great divisions of modernity, the grand ideologues and the believers in nothing. The rest of us have to make our way between them.

## The Fall

Once, Calasso says, there were rules about how things had to be done. There were rules about sacrifice, about worship, about planting and harvesting, about marriage, about kingship, even about war. Through repetition these rules become part of the very fabric of our consciousness, the way we imagined ourselves and the world. Then we threw them away. Our science taught us that they were false, not divinely ordained but invented by other men, perhaps for their own enrichment. Our political revolutions taught us that we could cast them all aside and be free.

And yet somehow it has not worked out as we hoped. Instead of a free and happy time Europe entered an era of tyranny and war, and the sentiment hung everywhere that were it not for the shreds of the unfree past we still maintained, things would have been even worse. How did it happen? In simple terms, the old regimes imposed limits on the rulers as well as the ruled, and when we swept them away we freed tyrants to dominate us far more cruelly than kings ever could. Yet the loss penetrated much deeper, into the structure of our minds. "Every obligation was a root," Calasso writes, and having cut away the chains that held us down we find ourselves cut off from what sustained us and too easily blown this way and that, from the extremes of devotion to Party and State to extreme indulgence and narcissism.

It is a remarkable picture, and one to which Calasso has added in many books since. But is it true? I find myself of two minds. One way to think about the twentieth century is to say that the cruelty is not new at all, just the technology; if medieval kings had had tanks and secret police forces they would have been just as awful. But if that is not true, if there is something about Hitler, Stalin and the Holocaust that is genuinely different from the woes of the past, then Calasso's analysis is the only one that to me has the breadth and power to explain the modern catastrophe.



# The Signal and the Noise: The Art and Science of Prediction by Nate Silver

*The Signal and the Noise: The Art and Science of Prediction* was written by Nate Silver, a consultant-briefly-turned-poker-player-turned-political-analyst who is most famous for the election forecasting website FiveThirtyEight.com. *The Signal and the Noise* is one of the small number of books – along with Philip Tetlock's *Superforecasting* – that aim to seriously assess the question of how predictable the future is, and how people can systematically improve their forecasting ability. This is a question which is of a lot of interest to me. Improving judgements about the future seems to be highly important in many areas (what will the effects of a policy be? When will different technological developments occur? How many people will die from COVID?) and very little attention is paid to it. I found *The Signal and the Noise* to be thoughtful, and I learned a lot from it.

## 1

A running theme of this book is that humans don't have very good track records predicting the outcomes of complex systems. But one domain where humans have excelled is weather forecasting. Weather forecasts are *amazingly* accurate relative to the complexity involved. In the mid-70s, the US National Weather Service was off by about 6 degrees (Fahrenheit) when trying to forecast three days in advance. This isn't much more accurate than what you get if you look at long-term averages – as in, what temperature is most likely in this region at this time of year, not taking into account any specific information. Now, the average miss is 3.5 degrees. This is actually slightly *less* of an improvement than I would have guessed, although to reduce the error in a forecast by a factor of two requires way more than twice as much effort, since errors can compound.

I was surprised to learn how large a role humans still play in weather forecasting. Having a human expert use their judgement in assessing many computer-generated forecasts is better than any of the forecasts are by themselves. Humans make precipitation forecasts 25% more accurate than computers alone and temperature forecasts 10% more accurate. Moreover, the accuracy added by humans has not significantly changed over time, so humans have been getting better at the same rate as the machines (!). If you're wondering why the weather forecasts you use don't *feel* very accurate, it's in part because weather services are private companies that tend to exaggerate forecasts for appeal; you won't see this inaccuracy in government forecasts. In particular, meteorologists are known to have a "wet bias" – they forecast rain more often than it actually occurs.

There have been some pretty tremendous positive social externalities of commercial weather forecasting, most notably in creating sophisticated early warning systems for extreme weather. The ability to predict typhoons in India and Bangladesh, for instance, has probably saved many thousands of lives. Silver has a few stories in here about people who refuse to leave their homes during an evacuation because of an

unjust scepticism of the forecasts. There also appears to be an exposure effect going on: studies of hurricanes find that having survived a hurricane before makes you *less* likely to evacuate future ones.

## 2

The terms ‘fox’ and ‘hedgehog’ used in this book come from the Greek poet Archilochus, who wrote that “a fox knows many things, but a hedgehog knows one big thing”. Foxes are people who don’t have grand unified theories, who constantly revise their beliefs to account for new evidence, and live in uncertainty. Hedgehogs are partisans, and have overarching worldviews which they’ll contort the evidence to fit.

The legendary psychologist Philip Tetlock ran a forecasting tournament in which he tracked and graded the predictions of political experts including professors and government officials over nearly two decades and which he summarised in his book *Expert Political Judgement*. The main finding: experts are barely more accurate at prediction than chance, and usually perform worse than simple extrapolation algorithms. There were too many hedgehogs and not enough foxes. The incentive for pundits and journalists is not to actually be accurate; it’s to appear reasonable while giving novel and entertaining predictions. Indeed, another of Tetlock’s major findings is that the *more* often an expert was on TV, the less accurate their predictions were.

Tetlock also found an overconfidence effect: when an expert says something has *no* chance of happening, it happens 15% of the time. When they said it is guaranteed to happen, it happens 75% of the time. While foxes get better at predicting with more information, hedgehogs get worse. If you have grand theories instead of partial explanations, having more facts can make your worldview even *less* accurate. Partisan differences in prediction were not seen in general (people were relatively unbiased in guessing how many seats republicans vs. democrats would win) but there were marked in specific cases (a left-leaning pundit is much more likely to say a *specific* democrat will win). These predictions were graded using a Briar score.

(I wonder if this generalises? If we have some kind of broad philosophical or political worldview that biases us, we might actually see *more* bias the more we zero in on specific cases. Hence, while talking about specifics and partial explanations is usually the better way to get at the truth, to be effective it might require some deconstructing of one’s prior beliefs.)

## 3

The woeful state of prediction might lead you to worry about climate science, where government policy is explicitly shaped by expert forecasts. Indeed, the magnitude of warming from climate change has been overestimated by scientists historically. The actual level of warming was below the 1990 IPCC estimates’ most optimistic projection. In response, the IPCC revised down its models in 1995, and now the observed outcomes fall well within the confidence interval of the projected outcomes (albeit the warming is still slightly less than predicted). You can certainly tell a story here about bias: scientists probably want to find a large warming effect and they think (correctly) that we’re at way more risk of panicking too little than too much. However, these estimates assumed a “business as usual” case; so, one factor that wasn’t addressed

adequately was that Chinese industry caused an increase in sulphur dioxide concentration starting around 2000, and sulphur dioxide causes a cooling effect. People forget about the other factors that contribute to warming – I was unaware that water vapour is actually the factor that contributes the most to the greenhouse effect! This all seems complicated to take into consideration so the less-than-stellar prediction performance of climate scientists can probably be forgiven. They also seem to have humility: just 19% of climate scientists think that climate science can do a good job of modelling sea-level rise 50 years from now, for instance. At least as of when this book was published (2012), the effect of climate change on most extreme weather events also appears to be unclear. This is a level of uncertainty that the media definitely fails to communicate.

Notably, the estimates around climate change are spectacularly noisy, which is well-known, but I think I had failed to appreciate just how noisy they are. Over the last 100 years, temperature declines in one quarter of decades – for instance, global temperatures fell from 2001 to 2011.

Another thing people seem to forget is for how long we've known about the greenhouse effect. It was discovered by Fourier (the Fourier transform guy) in the 1880s, and Arrhenius in 1897 was the first to predict that industrial activity would lead to a warming effect.

#### 4

The economist John Kenneth Galbraith famously said that “the only function of economic forecasting it to make astrology look respectable.” Indeed, at least in terms of asset pricing, we shouldn't expect economics to be of any help at all because of the efficient market hypothesis (EMH). This says that stocks and other financial products are priced in such a way that encapsulates the sum total of the information available to the market, such that individual trader advantage is rare. There are two components to EMH, which Richard Thaler calls the No Free Lunch assumption and the Price is Right assumption. No Free Lunch, or, colourfully, the Groucho Marx theorem, says that you shouldn't be willing to buy a stock from anyone willing to sell it to you; in other words, it's difficult if not impossible to consistently beat the market. The Price is Right says that assets are efficiently priced in a way that encapsulates all information.

Thaler has made a career out of exposing the extent to which economic models do not take sufficient account of human irrationality, and he is the ideological arch-nemesis of Eugene Fama, the father of EMH (they're also golfing partners, which I think is cute). Thaler has a famous paper in which he looks at the company 3Com, which created a separate stock offering for its subsidiary Palm. There was a scheme whereby 3Com stockholders were guaranteed to receive three shares in Palm for every two shares in 3Com that they held, which implied that it was mathematically impossible for Palm stock to trade at more than 2/3rds of the value of 3Com stock. Yet, for several months, Palm actually traded *higher* than 3Com, through a combination of hype and transaction costs. The final point that Silver makes about EMH is that it's in this fascinating epistemic state where if people actually believed it was true, it would stop being true. The only reason people trade stocks is because they think that *they* have better judgement than the market. This mirrors a lot of what people say about startups: if people actually believed

that almost every possible great company idea has already been taken, then they wouldn't start so many companies, undermining the process that made the original statement close to true.

Why does Silver talk about a theory of asset pricing so much? Because it's epistemically important to forecasting. If there's an efficient market for ways to improve the world, then if something were a good idea, someone would already be doing it. If there was an efficient market for ideas, every good idea would already have been tried and rise to the level of scientific consensus. And yet science is subject to massive systemic flaws, and huge opportunities for improving the world remain untapped because of inertia and apathy. Improving our forecasts of the future is important. It seems like a lot of people stand to make a lot of money from doing this. It seems like a small community mostly consisting of nerds on the internet would not be able to massively advance this field. But this impression is wrong.

Silver points out that if you look at the predictions of the Blue Chip Economic Survey and The Survey of Professional Forecasters, the former has some forecasters which do consistently better than others over the long run, but the latter doesn't. The reason why is that Blue Chip isn't anonymous, and so forecasters have an incentive to make bold claims that would garner them a lot of esteem if they turned out to be true. One study found a "rational bias" - the lesser the reputation of the institution that someone was forecasting from, the more bold they were in the claims they made. While considerations of esteem probably worsen forecasts overall, they lead some individuals to consistently outperform the crowd.

All of this should help us to understand bubbles. If EMH is true, how could outside observers notice massive market inefficiencies? Robert Shiller pointed out how the price-earnings ratio (share price divided by earnings per share) during the dot-com boom was unreasonably high, which was the sort of thing that had previously preceded a crash. One of the reasons why the bubble did not sort itself out despite people like Shiller pointing this out is the career incentives of traders: if you bet against the market and the market doesn't crash, you look like an idiot, while going along with the herd won't result in exceptionally bad personal outcomes. Silver says there is significant evidence that such herding behaviour exists.

Given all this volatility, it shocked me to learn that, over the long run, house prices in the US were remarkably stable until recently. In inflation-adjusted terms, \$10,000 invested in a home in 1896 would be worth just \$10,600 in 1996 (as measured by the Case-Schiller index). The value of such an investment would then almost double between 1996 and 2006!

## 5

There are a lot of interesting applications of the lessons from the science of prediction. One of the most exciting to me is predicting what research is going to replicate. One of the key lessons we should take from *The Signal and the Noise* is that academics, like everyone else, have all sorts of motivations, including prestige. Through honest motivations, scientists might go along with results that conform to their expectations and worldview, but that a financial market wouldn't price as being likely to actually be true. While markets have problems (see above), they're a vast improvement over hearsay and surveys. A 'prediction market' works because it actually incentivises people for

accuracy in a way they almost never are in other domains. It also works in part because of the wisdom of crowds: group aggregations of forecasts outperform individual ones by 15-20% on average.

Many of you will know this story: John Ioannidis publishes a paper with the provocative title *Why Most Published Research Findings Are False* which argues that due to the high number of researcher degrees of freedom, and the large variety of results that can be demonstrated with sophisticated statistics, most published research is probably wrong. More than a decade later, he seems to have been proven right. Bayer Labs found that more than 2/3rds of psychology research papers failed to replicate. Hence, the possible gain from a prediction market in study replication is large. One such project is [Replication Markets](#).

## 6

[Scott criticises](#) how people sometimes use the low total death tolls from terrorism as a way to mock conservatives, or people who are concerned about terrorism in general. Most years, lightning kills more people in the US than terrorism, but lightning also kills more people than have died from atomic bombs since WW2. The tail risks *are the whole goddamn point*.

If you've read *The Black Swan*, you'll know that lots of things are like this, with 'heavy-tailed' risk, and that we sometimes try to shoehorn these into Gaussian distributions.

Earthquakes are distributed according to a power law, whereby for every one point increase on the Richter scale, an earthquake is ten times less likely. So the bulk of the devastation comes from just a few earthquakes. The Chilean earthquake of 1960, the Alaskan earthquake of 1964, and the Great Sumatra Earthquake of 2004 accounted for half of all energy released by all earthquakes in the world over the entire 20th century! What else is less like height and more like earthquakes?

## 7

In one of the book's middle chapters, Silver uses terminology about infectious disease that many of us have become familiar with over the last couple of months, particularly SIR models. One nice titbit he talked about was the failure of SIR models to account for how there wasn't a re-emergence of HIV in the early 2000s among active gay communities like that in San Francisco (there was an increase in unprotected sex and other STDs). It's actually still somewhat a matter of debate why this happened but probably it was because people began to "serosort" – namely, choose partners who had the same HIV status as them. This goes against one of the major assumptions of the SIR model, which is that interactions among individuals are random.

The next few pages blew my mind the most out of anything I had read in a while. I can't believe I hadn't heard of President Ford's 1976 campaign to vaccinate 200 million people against a suspected H1N1 pandemic. The vaccine dramatically increased the rates of the nerve-damaging Guillain-Barré syndrome, and the public turned against it, such that only 50% of people were willing to be vaccinated! How has this not been brought up in the context of COVID?

**8**

I recommend this book, particularly if you're not already familiar with Philip Tetlock and forecasting. For people who are already interested in this kind of thing, I can still recommend skimming.

# The Silmarillion by J. R. R. Tolkien

I.

“In the beginning the Universe was created. This has made a lot of people very angry and been widely regarded as a bad move.” -- Douglas Adams

In 1755, an [earthquake](#) almost completely destroyed the city of Lisbon. Thousands of people were killed in the span of a few seconds by collapsing buildings, and thousands more died in the ensuing tsunami and fires. Aghast at the scale of the suffering, many intellectuals asked themselves: why did this have to happen?

In most human belief systems throughout history, this kind of question is easy to answer: someone pissed off the gods. This conveniently moves the conversation on to more manageable ground. The gods may have caused the earthquake, but it's a human's fault. Someone violated a taboo, failed to participate correctly in a communal ritual, or otherwise did something we don't like and we'll hold them responsible.

Many of us today hold our gods to higher standards. Whoever is responsible for the intricate reality around us must also bear responsibility for the suffering that takes place here, whether that's the God of Abraham or the designer of a simulation. The fancy theological word for this is "theodicy".

Ah, you may be saying, I don't believe in a bearded old man in the clouds or the geeky simulation hypothesis. The world just is, and that's that. Fair enough, but what of parents? They aren't responsible for the world, but when people choose to become parents they bring a new person into this world, a person who will certainly suffer during their life.

That brings us to J.R.R. Tolkien. He's famous as a novelist, not as a theologian or philosopher. Or to put it another way, he's famous for writing *The Hobbit* and *Lord of the Rings* and is most definitely *not* famous for writing *The Silmarillion*, a difficult book that immerses the reader in Tolkien's world unmediated by the modernist perspective of Hobbits. It variously reads like the Bible, an encyclopedia, and a short story collection. It's fiction, but it's definitely not a novel.

For reasons that can't detain us here, readers today almost exclusively consume written fiction in novel form, so it's only worth going on this limb if you fall into one of three categories:

1. You loved reading *Lord of the Rings* and you want to read more about the world of Middle-Earth no matter how dry the material. This is the largest category by far, but if you're in it, then statistically speaking, you have already read *The Silmarillion*, so this review won't further address you except to try to unpack something about the text you likely didn't dwell on while reading it.
2. You love reading ancient mythological texts and appreciate their depth of thought and archaic prose despite their textual contradictions, frustrating lacunae, and so on. This review may help explain why *The Silmarillion*, despite being an *imitation* ancient text, is nevertheless worth your time, but honestly it won't really do justice to what might appeal to you. Given the paucity of imitation ancient texts, you probably owe it to yourself to try it regardless.
3. You wonder why God would allow bad things to happen to good people, how a smart theist could possibly worship a God who permits suffering, or similar questions. Or maybe I have you wondering what a bunch of silly stories about Elves could possibly say about this subject. In that case, read on.

## II.

The fascinating textual history of the *Silmarillion* is probably known to a substantial minority of readers and would bore most of the remainder, so it will suffice to say that although Tolkien was already writing parts of it in 1914 at age 22, he worked on it off and on for his entire life and it remained not just unfinished but in a state of disrepair when he died in 1973 at age 81. It can be a frustrating read because the amount of detail varies wildly. Sometimes events or concepts are related in the barest summary, others are given the sort of blow-by-blow narrative one would expect from reading *Lord of the Rings*. The style likewise varies, but Tolkien most often attempts to mimic an archaic prose style, a style which, though familiar to him from his day job as a professor of Anglo-Saxon, is strange and alienating to many readers.

The text that is likely the most difficult for a modern reader to enjoy is the opening chapter, which describes the creation of the world by God and the fall of the mightiest of his angels into evil. I suspect many readers end up skimming this because of its fusty, quasi-Biblical tone. "Ugh, Tolkien was a conservative Catholic so I guess he thought this was necessary," some have probably said, turning the pages faster to find something more interesting.

Tolkien being so evidently conservative, it's easy to pass by too quickly and not recognize just how radical this chapter is. When Tolkien's friend C.S. Lewis wrote *The Chronicles of Narnia*, he took the standard Anglican theology of his day and just extended it into the world of Narnia. Aslan wasn't a metaphor for Jesus, he *was* Jesus, talking and behaving exactly as Jesus would (in Lewis' estimation).



Because it involves no characters from anything like our time, it's easy to forget Tolkien's conceit is that his Middle-Earth is *our* Earth, just in the past. His vision of the transcendent creator god, Eru Iluvator, isn't the god of another place, nor is it a Lewisian move where God takes a different form in a different land. It's merely the God of Abraham's name in the language that happened to be spoken in one part of the world in the ancient past.

In theory, *The Silmarillion* is still compatible with our book of Genesis. The first generations of humans are not described, so the Garden of Eden and the original sin of Adam and Eve can be assumed to happen concurrently with the Elves fighting their hopeless war to recover the Silmarils. But despite the fact that the Garden of Eden fits into the gaps of his own timeline, Tolkien's creation account differs sharply from that implied by Genesis and accepted by most Christians.

The Genesis account goes something like this: God created the world, and it was good, probably perfect. Then, in the Garden of Eden, humans listened to a literal lying snake and then sinned by eating a forbidden fruit. God was pretty upset about this and cursed men, women, the world around us, and even the snake.

So if you experience a devastating earthquake, why did it happen? Yep, someone pissed off God, but it's not your neighbor who skipped church last Sunday. It's no one alive, in fact. It turns out it was your extremely ancient ancestor who pissed off God, and no, God is still not over it.

Okay, it's not quite that simple. Many Christians read this part of Genesis metaphorically and in any case theologians have all sorts of different takes on this, but the point is that God created a good world, and later doctrine became very convinced that God could not possibly create evil. So most Christian theodicy involves maneuvers where God gave us free will, which was really good, but to be truly free, it had to be possible for us to choose evil, so therefore it's all our fault and none of His that our wrongdoing forced him to curse creation into occasionally killing us for no reason. What's that? Why was there a lying snake in a perfect paradise? There's very little in the Bible itself about this, but for example the Fourth Lateran Council decreed "the Devil and the other demons were created by God good in their nature but they by themselves have made themselves evil". The snake chose to lie, in other words, and hey, that's on the snake.

Back to Tolkien. In his creation story, God first creates angelic spirits who are later going to look suspiciously like a polytheist pantheon but who for now are there to assist him in creation. Rather than God speaking the world into existence, the world is sung into existence by this choir of angels, elaborating on a theme given to them by God. Alas, the greatest angel is proud and wants to sing a theme completely of his own

devising. A minority of angels take up his countervailing song and the result *seems* like a dissonant clash:

“And it seemed at last that there were two musics progressing at one time before the seat of Iluvatar, and they were utterly at variance. The one was deep and wide and beautiful, but slow and blended with an immeasurable sorrow, from which its beauty chiefly came. The other had now achieved a unity of its own; but it was loud, and vain, and endlessly repeated; and it had little harmony, but rather a clamorous unison as of many trumpets braying upon a few notes. And it essayed to drown the other music by the violence of its voice, but it seemed that its most triumphant notes were taken by the other and woven into its own solemn pattern.”

Then God makes a strong claim to Melkor (Satan):

“And thou, Melkor, shalt see that no theme may be played that hath not its uttermost source in me, nor can any alter the music in my despite. For he that attempteth this shall prove but mine instrument in the devising of things more wonderful, which he himself hath not imagined.”

I think most people zoom past this, but God is giving us a very hot take: “I created *everything*, even evil. I did it because evil is good, and it’s going to make the world is better than it would be without it.”

### III.

In 1710, Gottfried Leibnitz took a break from his feud with Isaac Newton and coined the term “best of all possible worlds” while writing about the problem of evil. Leibnitz was a genius polymath, but his theodicy is mainly remembered now for how hilariously Voltaire dunked on it with the “Professor Pangloss” character in *Candide*. Pangloss keeps insisting this is the best possible world while he’s reduced to poverty by war, loses an eye and an ear to a syphilis treatment, witnesses a friend drown during a shipwreck, sees Lisbon destroyed by, yes, the calamitous 1755 earthquake, and is tortured and sentenced to death by the Inquisition.

It’s not that no one has taken the “best of all possible worlds” stance since, but typically it’s accompanied by handwaving to distract from the pointless suffering of the world (often prompting aggressive atheists to go around rubbing theist noses in the world’s horrors). This handwaving is something that an author like Tolkien is well-positioned to

do. Tastes have grown darker as of late, but lots of traditional fantasy stories involve some excitement and moments of seeming danger, but then everything works out.

J.R.R. Tolkien fought in the trenches in World War I, and though he survived the destruction of his unit through extreme good fortune, all but one of his close friends from school died. He was also the sort of declinist conservative who has an unshakeable belief that the changes in the world around him are making it worse and worse. So no, he didn't believe that everything would just work out. Not only does he not wave aside suffering, he focuses on it, steel-manning the case against his theodicy even as he makes it. He has God announce that evil will make the world better in the first few pages of the *Silmarillion* and then, lest you think he wasn't being serious, the stories that follow are absolutely filled with suffering caused by evil.

## IV.

Consider Melkor, greatest of angels, who thinks to rebel against the theme of God's creation. What's his next move? Does he sidle up to an Elf in animal form, politely suggest they violate an arbitrary rule, and then basically disappear from the rest of the story?

No, he charges down to Middle-Earth and is so energetic in attempting to ruin creation that the world is preserved only through the efforts of a host of good angels. Multiple wars are fought and ultimately Melkor is captured and imprisoned for a time. After a while, the good angels release him since he claims to have learned his lesson. Needless to say, he has not.

By this point, there are no humans yet, but many Elves are now living side by side with angels in the "blessed lands" in the extreme west. The world is lit not by the sun and moon, which have not yet been created, but by the light from two beautiful shining trees. An Elf prince named Fëanor, the greatest craftsman who ever lived, captures the light from these trees in three magnificent jewels called Silmarils. Unfortunately, the supremely gifted Fëanor is also proud and possessive of his creations. The supposedly reformed Melkor deceives Fëanor into distrusting his brother, Fingolfin. When Fëanor threatens Fingolfin with violence, the good angels finally realize Melkor remains evil. Fëanor is exiled and the angels pursue Melkor, but they fail to capture him.

So far, this feels like it could be a more elaborate version of Genesis' lying snake.

But Fëanor's exile isn't permanent. The angels try to patch things up with a big celebration and Fëanor is uneasily reconciled with his brother. During the party, Melkor sneaks into paradise and poisons the two trees that light the world. The only way to restore them is to break apart the Silmarils, but Fëanor refuses to allow his creations to be destroyed even though he knows it's the only way to repair the only natural sources of light the Elves have ever known. Then it's discovered that on his way back out of paradise, Melkor stole the Silmarils and killed Fëanor's father.

This makes Fëanor, as one might expect, extremely angry. He blames the good angels for allowing this to happen on their watch, but as angry as he is at them, he knows the real culprit is Melkor himself. So he and his sons swear a terrible oath to use violence to recover the Silmarils from anyone who won't return them. Since that's currently Melkor, Fëanor calls on all his people to leave paradise and sail to the central continent of Middle-Earth where they will fight a war against Melkor to get the Silmarils back. Fëanor leads his most radical followers to the coast of paradise and demands the sea Elves there give up their ships. When they won't, Fëanor and his followers take them by force, killing all who stand in their way.

## V.

Obviously this is a story in close conversation with the Garden of Eden story in the Bible. But as with the creation story, this is not Narnian equivalence where the names are changed but the ideas are the same. Nor is it simple elaboration:

- Melkor doesn't just tell a single lie, he apparently tells lots of lies, and when they don't prove sufficient, he wounds the world and kills Fëanor's father.
- Fëanor doesn't get cursed and thrown out of paradise on a technicality or a momentary bad decision. He experiences loss and grief that, whatever his faults, he certainly didn't deserve. He demands that angels act to fix the evil in the world, and when they don't, he forges weapons, declares war on Satan, and personally leads an army out of paradise.
- From Tolkien's perspective, Fëanor's principal sin is pride: it's not his role to fix the evil in the world and not even remotely in his power. But pride is being kind of an abstract sort of sin, Fëanor almost immediately--but believably--commits the obviously unacceptable sin of murder, whereas in Genesis it's not until Adam and Eve's son Cain that someone does something indisputably evil and sheds blood.
- There's no room for fuzzy interpretations, like hey, Adam was duped by Eve and maybe women are the real villains here! Although some Elves are born later, since they don't age, most important characters in the rest of *The Silmarillion* are already alive and willingly choose to follow Fëanor.

- But not every Elf does! Some Elves, including some of Fëanor's people, do in fact realize this is a very bad idea and stay in paradise. This demonstrates the contingency of the choice made by Fëanor and the other rebelling Elves.

From all this, it seems clear to me that rather than transpose the Garden of Eden into a fantasy world, Tolkien sets out to improve upon it. Most of all, Tolkien's version foregrounds the problem of evil in a way that Genesis never does. Fëanor's crimes would never have happened if not for Melkor.

Who's fault is Melkor? The dominant Christian explanation (that God gave Melkor free will and, good golly, God definitely didn't want him to go and murder Fëanor's father, perish the thought) might actually apply, but only to the angels. They free Melkor from his imprisonment, restoring his agency, and then are dismayed when he uses his freedom for evil. But *The Silmarillion* is clear that although the angels are greater in every way than humans or Elves and know *something* of the future, they don't know all or even most of what will happen and they make mistakes. The narrative (ostensibly drawn from Elvish sources) is very respectful of the angels and doesn't actually say that freeing Melkor is a mistake, but then again it doesn't deny the obvious conclusion that it was. The only defense seems to be that these good angels are just so good they can't help but err on the side of mercy and they can't really understand evil.

There's no corresponding ambiguity in the story when it comes to God. Although Melkor thinks he's acting of his own accord, God was very clear in the first pages that because he created Melkor, everything Melkor does is what God desired to happen and the world is better for it. There's no free will, no mistakes of judgment, no erring on the side of mercy...the buck stops with God. He made this happen and he did it because it's good.

That brings us back to Voltaire's objection: how can this possibly be? The world is now covered in darkness, the sacred soil of paradise is stained by the blood of murdered Elves, and despite Fëanor's amazing talents, defeating Melkor is greatly beyond the ability of any Elf or human, and Fëanor's war will only cause even more suffering.

## VI.

"Out of life's school of war -- What does not kill me makes me stronger." --  
Friedrich Nietzsche

Maybe suffering, although not itself good, produces a good outcome by building character and teaching us valuable lessons. Nietzsche's formulation has an important qualifier that he needs to survive the lesson, but Christians who believe in an afterlife sometimes stretch this to include death as well.

Unlike Voltaire, Tolkien usually doesn't have characters give direct voice to philosophical arguments, but he offers an indirect answer to this point in the story of Túrin Turambar. This is one of the most detailed stories in the *Silmarillion*, and there was enough extant material that Christopher Tolkien eventually published a slightly different version as a separate (though very short) book. A lot of what happens to Túrin is taken from mythic precedents that most English-speaking readers aren't very familiar with (the Icelandic *Völsunga saga* and the Finnish *Kalevala*), but basically Túrin is a tragic hero like Oedipus, someone who does some great deeds but through divine meddling ends up both doing and experiencing really terrible things.

Tolkien's name for this story was "The Tale of the Children of Húrin", Húrin being Túrin's father, because it has an unusual frame which, as far as I know, has no mythic precedent. Húrin is a great human warrior who becomes a key ally of a secret city of Elves who are carrying on the war against Melkor even though Fëanor is long dead. Amid a surprise attack by Melkor's forces that routs the Elves and results in massive casualties, Húrin and his men end up fighting a heroic rear-guard action. Their efforts allow their Elvish allies to escape, but all of Húrin's men are killed and he himself is, at length, taken alive. Melkor wants to learn where the secret Elvish city is and he knows that Húrin knows, but Húrin won't say and resists even torture.

But Melkor is determined, and he is still the greatest of the angels. He puts a curse on Húrin's children, chains Húrin to a chair on a mountain, and uses magic to enable Húrin to watch for the next twenty-eight years as the curse turns everything his son and daughter do into suffering and they eventually die.

Does watching his children suffer make Húrin more noble? As usual, Tolkien's uncompromising narrative avoids the usual traps of trying to apply fiction to philosophy. For many of our heroes, suffering takes a form like: "Your own father cuts off your hand, which is quite painful, and you have no choice left but suicide". That's pretty bad! But then it's followed by: "Your suicide isn't successful thanks to a well-ventilated television antenna, and doctors give you an artificial hand that's apparently just as good as a real one". For Húrin, suffering means he watches helplessly as his beloved son and daughter's lives are ruined and they die in despair. There's no fixing this.

And, it turns out, there's no fixing Húrin. After his children are dead, Melkor simply lets him go, and Húrin spends the rest of his life as a bitter wanderer, taking some petty

revenge but along the way unintentionally laying the seeds for the destruction of the secret Elvish city he'd sacrificed his men's lives and resisted torture to protect. In case that wasn't enough, his actions also begin a process that will destroy yet another Elvish kingdom, one that wasn't even involved in Fëanor's rebellion and will be the last that stands against Melkor.

Now admittedly, Tolkien gives Húrin just a bit of an out here: it seems Húrin's understanding of his children's life weren't wholly true, seeing as his sole source was Satan News, and so when he's freed he has an overly negative interpretation of their lives and especially the actions of Elves toward them. Had I been Tolkien's editor I would have told him this really isn't necessary, Húrin's children's lives are extremely bad already without any deception, but whatever. The very understandable bitterness Húrin feels after being released is called "a madness" by the narrative, but after he does all sorts of damage to the free peoples of Middle-Earth, Húrin is finally "cured" by a Gandalf-like angelic being. So he's fixed? Well, no. Húrin is ashamed of all the trouble he's caused since being freed and commits suicide.

Whatever you make of the overcomplicated end of Húrin's life, I think the main takeaway is that when times were good, Húrin became a noble hero. Even when times stopped being so good, he sacrificed his life for others and resisted physical torture. But in the end, his prolonged suffering did not make him stronger, nor did he learn Important Life Lessons. Instead he was badly diminished and no longer capable of enjoying life even when it finally stopped being so bad. The Elvish pharmaceutical industry that produced such enviable products as lembas and athelas had no answer to his resulting mental health problems, and though he was eventually fortunate enough to be treated by a literal angel, her intervention fails and results in the loss of her patient.

All of this illustrates Tolkien's agreement with most people's common sense belief that while pain and suffering can sometimes help us learn, clearly it can also leaves us far worse off than we were before.

## VII.

Perhaps suffering really is bad, but maybe it's part of a larger plan. We have all done unpleasant things in service to a larger goal, and since God rules the entire universe, he may have some plan to take apparently meaningless suffering and cause something good to happen. The person suffering may not be the recipient of the good thing, and indeed may die never knowing that any good has come of it. But if the overall world

benefits from our suffering, that would be considerable consolation! And certainly good does come of suffering...sometimes.

Tolkien endorses a rather different version of this in *Lord of the Rings*, namely that great good can come of small good. The small mercy of Frodo toward Gollum, for instance, ends up saving the world from darkness in a way that Frodo never could have expected. So God sometimes uses small good to produce great good. Based on that, what do you suppose is produced from small suffering?

In case you can't guess, Húrin's story provides the answer. To the tragic hero Turin and his sister, most of the bad things that happen to them seemed like inexplicable bad luck. But Húrin (and we) are aware that all of this is actually part of a larger, supernatural plan operating in the world. Unfortunately, this isn't God's plan to turn all things to the good, it's Melkor's plan to turn the immense but ultimately individual-scale suffering of Turin and his sister into civilization-scale suffering. Even Húrin, after watching this play out in his children's lives, fails to understand his own part in Melkor's design and ends up doing the most to further it.

No one--Elves, humans, dwarves, or angels--can offer any genuine consolation for Húrin. What happened to him and his children was very bad and nothing can make that right. When angelic intervention allows Húrin the nearly unique opportunity to clearly understand the consequences of the suffering in his life and the lives of his children, his response is to completely despair and commit suicide. This is not endorsed by the text; Tolkien is always completely against despair. This means that Húrin's suicide is yet another way his suffering has diminished his heroism.

## VIII.

If suffering isn't a teacher and it's not part of a sovereign plan that will bring a greater good, why must Tolkien's characters suffer the way they do?

Here the story of Fëanor's brother Fingolfin is instructive. From the beginning, the narrative wants us to think Fingolfin is a good person and it does everything possible to distance him from Fëanor: the two brothers don't get along, Fingolfin tries honestly to reconcile but Fëanor is reluctant, and most importantly, Fingolfin isn't in the vanguard who resort to murder in order to steal ships. Although he's younger and seems to have none of Fëanor's artistic gifts, Fingolfin has far more followers because most Elves can tell he's the nicer guy. Beyond all that, Tolkien either didn't know or more likely just



didn't care about writing aphorisms like "show not tell" so he's willing to straight up say that Fingolfin is noble and good.

After sailing away from paradise with his relatively small group of supporters, Fëanor is too proud to want any help and burns his stolen ships rather than sending them back for Fingolfin's much larger force. This gives noble and good Fingolfin a perfect excuse to take an off ramp from this fiasco and stay in paradise like he should. But his people were fired up by Fëanor's speech, so rather than tell his people to do the right thing, he leads his army north and crosses the ocean over polar ice in a world that still is shrouded in unrelenting darkness, losing many of his followers to cold and crevasses. It's an impressive feat and one with a real element of nobility: where his brother harmed innocent people, Fingolfin and his people take the suffering on to themselves. Yet they are still doing this in support of Fëanor's war and so will suffer the consequences.

Lots more happens at that point, but suffice to say Fëanor dies, the angels create the sun and moon to give light, and Fingolfin ends up leading his dead brother's war. He lays siege to Melkor's stronghold for hundreds of years until, in a sudden attack, Melkor's forces break the siege and inflict a grievous defeat on the Elves. It's truly a catastrophe, but Fingolfin mistakenly believes the war has been completely and permanently lost. He gets on his horse and rides alone to Melkor's fortress where he bangs on the gate, challenges Melkor to a duel, and calls him a coward. Melkor finds himself unexpectedly forced to come out to avoid losing face in front of his forces. Although Melkor's evil has by this point greatly diminished him, he's still an angelic power, and however valiant, Fingolfin is just an Elf. After a long struggle, Melkor kills Fingolfin. Then a giant eagle who serves the good angels swoops down, scars Melkor's face with talons to distract him, and then carries off Fingolfin's body so it can be buried with honor.

*The Silmarillion* relates all this only in brief, so many read it without stopping to think: why didn't the eagle come and save Fingolfin's *life* rather than just his corpse?

This question echoes the most common complaint about the story of *Lord of the Rings*: "Why didn't they just put Frodo on one of those huge eagles and have it fly him straight to the mountain?" There are actually some pretty good reasons to think this wouldn't have been a good idea: Sauron has something of an air force, he's avidly looking for the One Ring and if he sees it coming on a huge eagle he will come out and go to the mountain himself to get it, etc.

But none of these objections apply to Fingolfin's case. If Melkor could have prevented the eagle from coming, he definitely would have done so and kept his face from getting scarred! Moreover, in the past when Melkor has behaved badly, the good angels always

go rushing out and then come back saying they can't find him. Yet here he is, outside his impenetrable fortress, fighting a long duel with Fingolfin. The angels must know he's there because they sent the eagle! Don't just send a bird, come yourselves and get him!

Why don't they? Why must Fingolfin be painfully killed? Why must his children grieve the loss of their father? Why is Melkor permitted to trudge back into his fortress to continue the endless war?

The answer is that it's a better story this way.

The despairing Fingolfin riding in wrath to the gates of Satan's fortress to challenge the devil himself to battle is a powerful moment, but most of its power comes from the fact that Fingolfin's charge is clearly hopeless. Again, for Tolkien, despair is always a mistake, and Fingolfin shouldn't have done this. Yet it's an extremely understandable mistake, and if the source of all the world's evil and suffering was sitting personified in an impenetrable arctic fortress, why, I myself would be tempted to--well, who am I kidding, I would hide under the couch, but if *someone else* strode forth to issue a hopeless challenge I would applaud their heroic spirit.

But the power of the moment depends on this being a hopeless, doomed action on Fingolfin's part. If an eagle saves him from suffering the consequences, instead of a courageous last stand we have a *Star Trek: The Next Generation* episode that safely resets at the end with no harm done. Fingolfin dying is a much better story than it would be if Fingolfin was saved. It's a better story than it would be if Melkor was permanently imprisoned by the angels and Fëanor spent his time figuring out how to make even more spectacularly glowing rocks in paradise. It's a better story than it would be if Melkor never fell and the world was nothing but sweetness and light.

The Silmarillion is a strange and deeply idiosyncratic work of fiction, but this is actually the most conventional aspect of it. Nearly all of our fiction involves some form of violence, danger, grief, loss, or death. This is because fiction that doesn't, that just depicts people chilling out and having good times, is--dare we say it--not as *good*.

Wait a minute. Was Tolkien's God *correct* to say that evil is good? And if so, was Voltaire wrong, and our earthquake-ridden, syphilitic world really might be the best of all possible worlds?

## IX.

Counterpoint: Elves aren't real. Fingolfin didn't truly suffer, because Fingolfin is made up. His name sounds goofy, not melodious like most of Tolkien's names, but regardless, he's fake. The suffering of fake people has no moral valence.

## X.

Why is it we are so immediately sure that fake people's suffering doesn't count?

In modern philosophy, there's a debate surrounding the "[paradox of fiction](#)" that asks why we have emotional responses to fictional characters even though we know they don't exist. There's no consensus answer, but the most common response seems to be something along the lines of: rationally, a human *shouldn't* feel emotion as we read about Fingolfin's desperate, doomed battle, but many people do, because humans are dumb (perhaps being coldly unmoved by even the most powerfully constructed fiction should be a new target for the rationalist quest for unbiased thinking).

This is common sense, but why? Well, Fingolfin is just a character in a book. He has no face, no voice, no thoughts.

Suppose we create a *Silmarillion* MMORPG and it has a Fingolfin non-player character who stands around Mithrim with an exclamation mark hovering over his head. The player can converse with Fingolfin using branching dialogue trees, and Fingolfin offers some otherwise unprompted remarks about gestures, fighting, or other events within a certain distance of the character. This version of Fingolfin has a face and a voice actor provides him with a voice. It's debatable what relationship dialogue trees and some scripted reactions have with "thinking" but certainly no one will get charged with murder when the game company goes out of business and shuts down the servers.

But suppose the game company doesn't go out of business, and it upgrades Fingolfin with the latest in gaming technology: now instead of simple dialogue trees, his dialogue and actions are determined by a neural network that has been trained on data that greatly elaborates on the minimal personal history Tolkien writes for Fingolfin in *The Silmarillion*. Does harming *this* version of Fingolfin have moral valence? I think the average person off the street would still say no, but most people reading this will review will probably frown and start asking questions about what sort of neural network it is, how many simulated neurons there are, and so forth. In other words, at some point we

think Fingolfin stops being a mindless puppet and becomes a real person whose experience has moral weight.

This happens when Fingolfin acquires enough attributes that he becomes, roughly, the same thing as we are. There's lots of disagreement over the exact line (e.g. which side are [chickens](#) on?) but that seems like the line we're drawing. Until then, he's just a representation of a person, not a real person. An image, you might say.

## XI.

"So God created mankind in His own image..." -- Genesis 1:27

Compared to the entity that created the universe we see all around us, which version of Fingolfin are we most like? The AI that runs cognitive processes of roughly the same complexity, just in a different medium? The dumb puppet that mechanically follows its script and has no faculties that could allow it to understand what it really is and why it does the things it does? Or the fictional character who has no independent existence at all except in the mind of his observer?

## XII.

But I should say, if asked, the tale is not really about Power and Dominion: that only sets the wheels going; it is about Death and the desire for deathlessness. Which is hardly more than to say it is a tale written by a Man! -- J.R.R. Tolkien

Tolkien wrote this about *Lord of the Rings*, and encountering the quote as a teenager I was nonplussed. *Lord of the Rings* seemed like it was about, well, not power and dominion, but maybe courage and faith and friendship. But death? Deathlessness?

It's possible to work this out with just the text of *Lord of the Rings*, but *The Silmarillion* makes it much easier to understand. To readers of *Lord of the Rings* (and certainly to those watching the movies), it's easy to come away with the impression that Tolkien was trying to say that the Elves were better than humans, that humans of the past were better than humans of the present, that in fact the past is better than the present in every way, and that it's bad that, say, the beauty of Lothlorien is ended when the One Ring is destroyed.

People get this impression because it really is close to Tolkien's own feelings. He really did feel like the world was getting worse all the time and he shared the Elves' love of the natural world. But the wise Elves of *Lord of the Rings* are nevertheless people who rebelled against paradise (Galadriel followed Fingolfin across the ice, while Elrond is on one side a descendent of Fingolfin himself), and impressive as Rivendell and Lothlorien seem, they're actually the pitiful remnants of the great war Fëanor launched.

In other words, Tolkien couldn't help agreeing with the Elves, but he nevertheless wrote that the viewpoint he and they shared was wrong. The Elvish rings might not have been tainted by Sauron, but in holding back decay, they hold back God's intended plan for the world. When the Elvish rings fail and their bearers finally sail into the west, it's their final admission of futility. Melkor was banished from the world, but only through the intervention of angelic powers, and although the Elves that remain helped to defeat his lieutenant Sauron, the world remains permanently afflicted with evil and everything they tried so hard to preserve has been lost. "We have fought the long defeat," is how Galadriel sums up her extremely long life to Frodo.

In college, I tried to explain to a friend Tolkien's admiration for "northern courage" (so-called because Tolkien associated it with Scandanavian culture), the act of fighting on even though the cause is hopeless. "Fighting when you're sure to lose sounds stupid to me," my friend said, and unfortunately it was only hours later that I realized what I should have said in response: our lives are a long defeat! Death is inevitable, yet we urge each other to struggle onward in the face of suffering, grief, and loss.

### XIII.

In Genesis, death is a curse, part of God's punishment for humans eating the forbidden fruit. But in another of his radical revisions to Christian doctrine, Tolkien recasts death as a gift given by God to humans. Alone among all the different kinds of beings in the world, when humans die their souls leave the world. Where humans go after this is completely unknown, but everyone else in the world--Elves, dwarves, ents, eagles, even the angelic powers--are "bound" to it until it ends. So although Elves can be killed, they don't age, or perhaps more accurately, they age at the same rate as the world does. For example, Galadriel in *Lord of the Rings* is approximately seven thousand years old and since she eventually sails west, it's reasonable to expect she will continue living until the end of the world.

To most reading this, it will sound like Elves got the better end of the deal. But unlike us, Tolkien was deeply committed to a declinist worldview. The world isn't as good as it

used to be, it's getting worse, and the Elves are trapped in it. For Elves, even death isn't an escape from their prison sentence: Tolkien never worked out the details, but he decided that when killed, Elves must eventually reincarnate and continue to share the fate of the world.

So death is an escape from the prison of an ever-shoddier universe, and the only reason humans fear this mighty gift is because Melkor lied to them and fooled them into desiring the "immortality" of being trapped in the world. This is why, in *Lord of the Rings*, the evil One Ring extends life, and this property is considered bad even by someone like Bilbo who benefits from it.

That is what *The Silmarillion* says explicitly about human death, but this same idea is present in the negative portrayal of Fëanor's desire to preserve his Silmarils. He doesn't want to let go of the beauty he personally made in the world, but the Silmarils are taken from him anyway. The Elves he leads away from paradise do everything they can to recreate what they liked about the paradise they left, but it's always just a pale echo of the real thing, and eventually even their lesser recreations are destroyed by Melkor and his servants. The entire war against Melkor, a war that goes progressively worse the longer it goes, is yet another instance of this theme.

Most readers probably won't share Tolkien's deep conviction that the world is getting worse, but when applied to the human body it cannot be denied. No matter who you are or what your particular health situation is, each moment is bringing you closer to decline and death. Whatever hopes some might have that in the next few decades technology may begin to change the human relationship with aging, at least we can agree that for the entire time humans have lived up to the present, the inevitability of death has been a central part of the human condition.

This is why Tolkien implies he can't help but write a story about death and the pursuit of deathlessness. He insisted his work was not allegory but "applicable", and certainly nearly everything he wrote about--the decay of the world of Middle-Earth, the Elves and their doomed war, the humans afraid of the "gift" God has given them--is readily applied to the human condition.

Why do Tolkien's characters suffer? Because they are part of *The Silmarillion*, a haunting work that, as its beginning promises, is "deep and wide and beautiful, but slow and blended with an immeasurable sorrow, from which its beauty chiefly came". His characters suffer because *we* suffer, and their lives are meant to relate to ours.

But if Tolkien's stories are applicable to the human condition, is his stories' theodicy applicable to our own universe? Decline and death is the *human* condition, but humans

are just one species out of millions, living on one planet that orbits one unassuming star out of uncountable billions. Even if we accept Tolkien's analogy that God is a writer, isn't it presumptuous to assume we're central to the divine narrative?

Before answering this question, a last point to emphasize about *The Silmarillion*: Tolkien chose music as a metaphor for creation because of the way different notes harmonize together, and accordingly his theme of death and the pursuit of deathlessness resounds in every part of his world, not just in the experience of his fictional humans.

## XIV.

When I was young, I read in a children's astronomy book that the universe is expanding but that astronomers aren't sure of its eventual fate. Maybe it will collapse back on itself in a "big crunch" that could in turn result in an endless cycle of big bangs and crunches, the book said, or maybe it will expand forever. For years, I felt extremely confident that it would collapse back on itself. It just seemed so much more elegant and tidy for the universe to exist in an eternally recurrent cycle.

One hesitates to make firm pronouncements on cosmology from our humble, one-planet civilization, but today the evidence seems very strong that the universe's expansion is accelerating and on an inescapable course to some variation of grim "heat death" where all matter and energy are frozen into permanent stasis. When I first learned this, I was mildly annoyed the universe was failing to align with what I personally thought was the preferable aesthetic outcome.

Having thought about this now and then since, heat death has started to seem much more harmonious than I first realized. After all, for all their inconceivable size, stars and even galaxies get old and die. And though inconceivably small, atoms and even particles eventually decay and die as well.

You might object here that the word "death" is being used too metaphorically, and that what happens to humans doesn't really have anything in common with the ends of atoms and stars. Maybe. But even so, it seems to me that "decay" is harder to argue. Entropy always increases. Tolkien was hopefully wrong about human society getting continually worse, but in a long enough run he was right about the world. The universe is decaying, and though we may have space-traveling descendents who build some science fiction equivalent of the Elvish rings that hold back decay, it may be that the best fate we can hope for humanity is the fate of the Elves, bound to the decaying universe until the end.

In such a universe, are we so sure that Tolkien's theme of death is just a parochial human idea? Are we so sure that the grief and loss you and I will suffer in this life isn't part of a harmony sung by the entire universe, from the tiniest particle to the largest galaxies?



# The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron by Bethany McLean and Peter Elkind

*"If I had to describe my sixteen years of corporate work with one phrase, it would be 'pretending to add value.'"*

—Scott Adams, creator of *Dilbert*

I.

We all have our failures, our embarrassing moments, our simian exigencies that ever urge us to short-term expedience. *Confiteor, mea maxima culpa, dimitte mihi*—all that good stuff, let me get back to practical matters, thank you very much.

But can you imagine living an existence so abject, so corrupt, so base, an enormity so egregious that you can't but burst and collapse into a veritable black hole that absorbs and annihilates everything in the immediate vicinity, such that in your self-contained universe you become a veritable byword and your story mandatory reading for anyone interested in pursuing a career even remotely related to yours?

*Ecce, Enron.*

Just as the sinking of the *Titanic* was for the 20th century, if not a pivotal moment, then at least an iconic one, I suspected the collapse of Enron will, along with the 9/11 attacks which precipitated the bankruptcy thereof in some complex economic way, be considered critical and definitive for the 21st century, because, so I was told, it revealed deep weaknesses in not just accounting controls (Enron's auditor, Arthur Andersen, was brought down with the ship) but also in the capitalist system itself. Enron, as a byword for corporate, has implanted itself in the public consciousness---hence my interest in this topic.

I wasn't terribly inclined to start such a massive book on such a sordid topic as corporate politics. But whenever I looked at the Wikipedia page for the Enron scandal with a mind to make some sense of the matter, every time I got to the section that details the causes thereof, I would read something like "Enron earned profits by providing services such as wholesale trading and risk management"—and just a few sentences later there I would hit the limit of my intelligence and education. I chose to read *Smartest Guys in the Room*, hopeful that in it I would find an accessible narrative; after all, in theory, one can only achieve "best-selling" status if people from diverse backgrounds can buy and read it.

And, after going through years of journalism, all distilled into nearly five hundred pages, the facts rendered as thoroughly and limpidly as they could ever be for a layman—I still don't really know how Enron made money in the first place.

Perhaps they never did.

*The Smartest Guys in the Room* is written in a very plain, unadorned style suitable for an investigative report. Its length, I suppose, is quite justified given the history of the company, the characters and their history, and the diverse endeavors (which, whether in accounting or business were largely, abstrusely fraudulent, or whose ethical considerations were, at the very least, questionable). This book seems largely written for documentary and historical aims, rather than to provide any esthetic edification to the reader—a work whose longeurs may be worthy for the annals but which is a bit tedious as a medium for a cautionary tale.

It wasn't too long into it that I gave up trying to understand what Enron was doing, because I realized that even attempting to do so might lead me into the same iniquity that had taken the corporation itself. If Enron's accounting maneuvers appear confusing, that was the point—they made them deliberately complex in order to hide their failures (and I deliberately use the word "failures" instead of "debts," because the latter has more accounting connotations, which I, for obvious reasons, want to avoid), to the point that it required years of hearings, investigations, and trials to figure out what the hell had happened.

## II.

The purpose of this section is to describe the substance of Enron's fraud in the plainest language I can muster. Perhaps this review can serve to help those who should find themselves in the same position as I was prior to my reading this book: wanting to know the details of quite possibly the biggest corporate fraud in history but getting continually caught up in the details and complexity.

The source of money—the location of money—the flux of money. If these three be the Trinity, then ACCOUNTING is perhaps the LORD thy God, in which the three are consubstantial. Come tax time, all of us too painfully know the moral obligation to track and plan our finances—and the punishments for not doing so.

A naive view of accounting—one even I admittedly held for too long, which perhaps bewrays the lack of complexity, and according meagerness, of my finances—would be something like: "What's the big deal? Money is money—you use it to buy things." But then that raises the question—what, exactly, is *money*? Is it just a store of value? But if that's the case, an asset like a computer, which is indisputably valuable, should be considered "money"—and, if that's the case, shouldn't fluctuations in its value (whether from upgrades or depreciation) be reflected somehow on the financial statements?

Because accounting can get so complicated so fast, and because each accounting maneuver is fertile ground for corruption, there are certain rules, laws, and audit procedures public corporations have to follow when it comes to their financial reporting. But, of course, with all rulesets, especially ones that are broad and complex, there's room for interpretation. Sometimes one person will think one thing (almost always to his benefit), another person (the person who loses) disagrees, there's a lawsuit, they go to the courts, i.e., capitalism.

In fact, there are even scholars who think ["objectivity in accounting is largely a myth . . . arguing that accounting should be approached as a form of 'dialogue' through which accountants can construct, 'read' and probe situations in a variety of ways."](#)

So, very briefly: in its years of business Enron took broader and broader interpretations of accounting rules and made transaction after transaction that had apparently no purpose but to

obfuscate their commercial failures—till their books had no connection to the reality of the corporation's economic position.

For example, Enron decided—which seemed quite innocuous at the time, as this was what many companies do—that they would declare the entirety of the profits of any new deal they signed *immediately* upon their signing it (a convention known as mark-to-market accounting). Dispensing with the inconvenient question of how could one possibly know how much a 10-year contract would return in the end—maybe something changes, maybe your partner doesn't want to do business anymore, maybe some commodity you depend on becomes prohibitively expensive due to factors outside your control—their financial statements used incomprehensibly complex and generous models or simply vague statements such as "Management's best estimates."

. . . the idea was all and the idea, therefore, should be the thing that was rewarded. [Jeff Skilling, the CEO,] felt that a business should be able to declare profits at the moment of the creative act that would earn those profits. Otherwise businessmen were mere coupon clippers, reaping the benefit of innovation that had been devised in the past by other, greater men. Taken to its absurd extreme, this line of thinking suggests that General Motors should book all the future profits of a new model automobile at the moment the car is designed, long before a single vehicle rolls off the assembly line to be sold to customers. Over time this radical notion of value came to define the way Enron presented itself to the world, justifying the booking of millions in profits on a business before it had generated a penny in actual revenues. In Skilling's head, the idea, the vision, not the mundane reality, was always the critical thing.

—Excerpt from *The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron*, by Bethany McLean and Peter Elkind

And here is one of the best advantages of doing business this way: say some deal falls through or you lose a bunch of cash, you can simply "revisit" those old models, change some assumptions, and thereby you can reveal profits you didn't even know existed!

Enron employed other tricks. Deal makers regularly revisited large existing contracts—some more than five years old—to see if they could somehow squeeze out a few million more in earnings. Sometimes the contracts were restructured or renegotiated; other times, they were simply reinterpreted in ways that made them appear more profitable. "When the last-minute call for earnings went out," says one high-level deal maker, "I'd go: 'Which contracts did I do five years ago that had potential value?' A lot of them you could remark." Skilling himself labeled the contract portfolio "a gold mine."

Earnings projections on mark-to-market deals, based on complex models, were reexamined. Was it possible to be a little more optimistic? A small move in a long-term pricing curve could generate millions in extra accounting profits. The curves often went so far into the future that drawing them was already little more than an educated guess. The danger was that skewing curves to generate more profits was not only improper but also raised the likelihood that the curves would turn out to be way off base, producing a big mismatch between Enron's projections and a reality it would eventually have to face. But that, of course, was a future concern, far removed from the crisis of the immediate quarter.

But these financial "disclosures" were deliberately engineered to be impenetrably complex. Before long, Bethany McLean, one of the co-authors of this book, revealing the nakedness of the Enronperer, wrote, in March of 2001, less than a year before the bankruptcy, an article for *Fortune* entitled *Is Enron overpriced?* which title, I am told, is still reckoned as the favorite for Most Prescient Understatement of the Century. And a courageous piece of journalism it was, to ask the blatant question that the fake news (NO STOP THIS IS NOT THE TIME OR PLACE) everyone was being willfully blind to and bear the risk of ignominy ("What, you don't understand Enron, the coolest, fastest-growing company in America? What the hell kind of business journalist are you?")

The factors that acted on Enron to crash its stock in the period of a month are many and complex, and I am not competent to summarize, or even understand, the minutia and their actors. But I hope, by my expenditure of time in reading this book, and this subsequent review, I can distill the essence of the collapse and thereby perhaps divert those who are too besieged by mundane concerns from perpetuating, whether deliberately or through willful blindness, a similar enormity.

When, exactly, did Enron cross the line? Even now, after all . . . that's an impossible question to answer. There have been accounting frauds over the years . . . [but in] such cases, someone at a company ha[d] to consciously consider the fact that he or she is about to commit a crime—and then commit it.

But for the most part, the Enron scandal wasn't like that. The Enron scandal grew out of a steady accumulation of habits and values and actions that began years before and finally spiraled out of control. When Enron expanded the use of mark-to-market accounting to all sorts of transactions—was that when it first crossed the line? How about when it set up its first off-balance-sheet partnerships . . . [or] when it categorized certain unusual gains as recurring? Or when it created EPP, that "independent" company to which Enron sold stakes in its international assets and posted the resulting gains to its bottom line?

In each case, you could argue that the effect of the move was to disguise, to one degree or another, Enron's underlying economics. But you could also argue that they were perfectly legal, even above board. Didn't all the big trading companies on Wall Street use mark-to-market accounting? Weren't lots of companies moving debt off the balance sheet? Didn't many companies lump onetime gains into recurring earnings? The answer, of course, was yes.

And my studies have concluded that the fall of Enron can be encapsulated with a single phrase: the inevitable fate of one who pursues the appearance of success rather than success itself.

### III.

*"Every single deal that I did at Enron when I was CFO was approved by Enron's accountants, the outside auditors at Arthur Andersen, Enron's attorneys, Enron's outside attorneys, the banks' attorneys, and Enron's board of directors."*

—Andy Fastow, CFO of Enron (from his corporate culture talk)

The biggest takeaway is that the Enron fraud was systematic and institutional, but not a *conspiracy* (N.B: I am using the word "conspiracy" as a layman. Some of Enron's officers were

convicted of the criminal charge known as "conspiracy," but I don't know what the technical standards/procedures are for inferring such a charge in the jurisdiction in which they were convicted.)

Systematic and institutional, in the sense that everyone on all levels, to varying degrees of cognizance thereof, participated in perpetuating the fraud (as the above quote of Fastow's illustrates); but not a conspiracy, in the sense that it seems there was never a group of people that, sitting down with the books, whispered: "We know we must do x under the law, but we're instead going to fraudulently and illegally do y because the rewards are high and maybe no one will notice, and we all agree to keep this a huge secret."

Rather, Enron's crime was born in its ethos, which can be articulated thus: "The way we do things—business, accounting, whatever—is right because we're Enron."

Smith was struck by things that most Enron employees had long since taken for granted. The abrasive, cutthroat culture. The condescension toward anyone who didn't work at Enron. And always, the obsession with the stock price at every level of the company. "Everywhere you looked, the stock ticker was going," said Smith. "In the lobby of the building. In the lobby on your floor. It was on the screen of your computer. Everybody was focused on the stock price. You couldn't get away from it. When the stock wasn't doing well, the mood changed."

A corporation having a toxic, cutthroat culture is hardly noteworthy. What is more extraordinary is that all the usual checks and balances in the capitalist system failed—from the outside auditors at Arthur Andersen who approved transactions despite that the email record shows they tried to raise concerns initially (partly because they were afraid of losing their biggest client), to the business analysts who praised the company, driving the stock up and up while the company produced nothing of value. Meanwhile, inside jokes about the "counterintuitive" accounting maneuvers openly proliferated—which, like SS's adoption of a skull as their insignia, makes you wonder how nobody realized that they were doing something wrong. (Probably anybody who did was fired.)

At the Congressional hearings into the bankruptcy, Andy Fastow (CFO) pled the Fifth, which anyone, even if the full extent of his legal education were a mere high school course in law, would have advised. Jeff Skilling's (CEO) lawyers thus advised him too, [and he ignored them](#). There are only two ways I can conceive of this: Perhaps Skilling was simply naive; but, given the description in the book of how he maneuvered his way up from a blue-collar family to become the CEO of one of America's biggest companies, I think this is unlikely. The better explanation is that this is a man so caught up in the cult of Enron that he truly believed (and, having gotten out of prison a few years ago, apparently still believes) the collapse was the result of pure politics, and that he and Enron had done nothing wrong.

. . . the very picture of a sophisticated, booming business: a big open room, bustling with people, all busily working the telephones and hunched over computer terminals, seemingly cutting deals and trading energy. Giant plasma screens displayed electronic maps, which could show the sites of [Enron Energy Service]'s many contracts and prospects. Commodity prices danced across an electronic ticker. "It was impressive," recalls analyst John Olson, who, at the time, covered the company for Merrill Lynch. "It was a veritable beehive of activity."

It was also a veritable sham. The war room had been rapidly fitted out explicitly to impress the analysts. Though EES was then just gearing up, [they] had staged it all to convince their visitors that things were already hopping. On the day the analysts arrived, the room was filled with Enron employees. Many of them, though, didn't even work on the sixth floor. They were secretaries, EES staff from other locations, and non-EES employees who had been drafted for the occasion and coached on the importance of appearing busy. One, an administrative assistant named Kim Garcia, recalls being told to bring her personal photos to make it look as if she actually worked at the desk where she was sitting; she spent most of the time talking to her girlfriends on the phone. After getting the all-clear signal, Garcia packed up her belongings and returned to her real desk on the ninth floor. The analysts had no clue they'd been hoodwinked.

Just as Enron's financial executives convinced themselves that their financial shenanigans stayed within the rules, it seemed, so EES executives reasoned that this deception wasn't a problem. Eventually, EES really would use all that space. Eventually, there would be hundreds of busy employees working the phones and trading energy and the division would be every bit as fabulous as they were telling investors. It just wasn't quite there yet. In many ways, Skilling's little Potemkin Village stood as the perfect metaphor for EES: so much of what outsiders were led to believe about the operation was at odds with what was really going on.

Enron is the apotheosis of the proverb: "Fake it till you make it."

## IV.

. . . it's hard not to wish for that naïve time when Enron was shocking, when we believed President Bush when he said that Sarbanes-Oxley would rein in greed, and when we really, truly thought that the act of putting Jeff Skilling and Ken Lay behind bars would solve everything.

At this point, if any should say that Enron is what you get when a society underregulates, that we need more oversight, regulation, revolutions, socialism, communism, etc., my response would be that they do not understand the nature of Enron. Legislation, regulation, or oversight is, at best, addressing the symptoms, and not the disease, as the Enron story perfectly illustrates.

*"Accounting rules and regulations and securities laws and regulation are vague. They're complex. . . . What I did at Enron and what we tended to do as a company [was] to view that complexity, that vagueness . . . not as a problem, but as an opportunity. The only question was 'do the rules allow it—or do the rules allow an interpretation that will allow it?'"*

—Andy Fastow, quote from *The confessions of Andy Fastow*, July 1 2013, *Fortune*, by Peter Elkind

Do you really think that your tiny little roadblocks will stymie, even slightly, the infernally ingenious engines of the capitalist machine?

In his *Meditations on Moloch*, Scott Alexander observed: ". . . many of the most important competitions / optimization processes in modern civilization are optimizing for human values. You win at capitalism partly by satisfying customers' values. You win at democracy partly by

satisfying voters' values. . . . But it's important to remember exactly how fragile this beneficial equilibrium is."

Yet Enron did not only find its way around this equilibrium—it prided itself in so doing. *Smartest Guys* documents Enron's involvement with the California electricity crisis of 2000, during which the state government, in an attempt to control the crisis, created a series of regulations; and Enron, being Enron, found loopholes in the complexity and exploited them for monetary gain, which often involved withholding power to the starved state, i.e., doing literally the exact opposite of satisfying its customers' wishes. This is just one of the many examples where Enron asked "Is there an opportunity for profit; and, if so, is it legal?" and, by a great contortion of imagination and lawyers finding that the answer to both was yes, went all out, regardless of the human lives affected as externalities.

My point is not that more people should have had harsher punishments. In fact, I'm not entirely sure what my point is. I guess what I'm saying is that it seems the problems with Enron run much deeper than questions of economic/social policy.

What about "being a good person"? Well, that's an even harder question than "what exactly happened with Enron."

V.

But if there's anything that's become clear in the last ten years, it's that what can go wrong with business is something that can't be isolated to Enron—or fixed by the simple act of making one man our pariah. In the end, that's why the Enron story will always be relevant: It's a tale of human nature, a morality play for our age.

*"When I became CFO, never once did I ask myself the question: was what I was doing ethical? . . . The only question I asked myself is: am I following the rules? And, in fact, I entered into transactions that I knew were intentionally, materially misleading. The rationalization I used was, hey, I'm following the rules; so it's okay. . . . If I don't do it, someone else will do it first, and they'll get an advantage."*

—Andy Fastow, during an interview on day 1 of the 2017 PLUS D&O Symposium

The whole Enron story got me thinking about the nature of lying and deception. Specifically, is a lie "wrong" if it doesn't technically break the rules?

Upon hearing the very word "lie" itself, the mind instantly responds with aversion, ascribing to the sound all these negative connotations. Why is that?

The religious would say that lying is enjoined by the commandments; the secular would say that it erodes social trust.

I find both answers unsatisfying. After all, there are numerous examples in the world's religious texts (and real life) where everyone understands that, in some specified circumstance, lying is the "godly" option; and we also all know that there is no correlation—and even an inverse correlation—between a person's truthfulness and his success in obtaining public offices (and my sister, a forensic accountant, assures me that for every one Enron, there are ten others that no one has noticed).

Furthermore, lying seems to be a prerequisite for doing any kind of business. I can't tell you how many times I've been told by people to "be honest" in job interviews—while simultaneously instructed to reframe my "weaknesses" as "strengths," to not say that you left your last job because management was toxic, to say that you are interested in this company in this particular rather than that you saw the ad for the position opening and you needed a job, etc.

But I see no meaningful difference in [not technically lying](#) and "actual lying."

Whom, exactly, was Enron trying to convince? Shareholders? Themselves? Someone else?

## VI.

The scariest thing is that I have absolutely no difficulty seeing myself perpetuating something horrific like Enron.

I'm a hypothetical Enron officer. I find all my worldly needs satisfied; I have endless money, women, respect, and influence. Ethics? Nothing but a word for ivory-tower philosophers; I'm focused on real-world *business*. Besides, I have too much to lose. Why should I speak up and thereby risk being destroyed by the powers which, only by great political dexterity, I exploit while simultaneously keep at bay lest they get out of control and destroy me myself?

Or, which is more likely to be the case for most people: I have a family I have to feed. My boss comes to me and says: "Approve this transaction which we are prepared to defend as legal with the full force of our lawyers, or you're fired." Should I really let petty workplace disagreements get in the way of providing for my loved ones?

So I did: I lied in the job interview. I lied about the progress of my work. I lied that it wasn't my fault. And why not? After all, everyone does it, and I don't want to be left behind.

And without fail, every single time, whether in a day or in a few years, somehow, somewhere, it has come back to punish me.

Am I, like Enron, just "bad at business"? Or is it that lies, even the not-technically-ones, scale very poorly; and we, as the myopic, short-lived apes we are, simply mistake superficial expedience—which, by its tenuous nature, cannot long abide—as good sense, and, when the edifice inevitably crumbles, are unable to trace the chain of ever-compounding falsities?

I hoped that writing this review would help me answer the questions that have tormented me my entire adult life; but, unfortunately, I fear that I'm nowhere closer. My only solace is that when I look back at all the times I didn't do my best, spent more time appearing productive than actually being productive, constructing political and legal devices rather than engineering ones, blaming others so as to divert responsibility from my position, to continue to collect a salary a not-insignificant multiple of the median for the population, and when I see all the judgments whose memories still, to this day, rise to the present moment and smite me once more, I can remember *Smartest Guys*, and maybe practice some self-forgiveness—for, though I may have been duplicitous, wasteful, and unproductive, at least I wasn't as duplicitous, wasteful, and unproductive as Enron.

But, behold: I see Enron's fixation on nice-looking financial statements and its stock price while the company essentially had no, even negative, productivity; I see some affinity between Enron's "not-technically-lies" and those in my own life; I see the times when I sought the mere



*appearance* of wealth, truth, etc. while neglecting them in *essence*. And though I still absolutely cannot understand, much less apply, that strange and incomprehensible apothegm "thou shalt not bear false witness," I do think I'm beginning to have some glimpses as to why "Thou shalt not make unto thee any graven image" is canonically superordinate.

# The Sovereign Individual by William Rees-Mogg and James Dale Davidson

The Sovereign Individual, aka TSI, aka the most important book you've never heard of, aka the book that predicted Brexit, is a history book about the future.

The authors, William Rees-Mogg and James Dale Davidson (WRM & JDD) admit it is tough to predict the future, and quote Mercedes saying in 1903 "there would never be as many as 1 million automobiles worldwide. The reason was that it would be implausible that as many as 1 million artisans worldwide would be trainable as chauffeurs". But they're going to try anyway.

They have a good response to critics who think predictions can never be worth anything: if you drop a \$100 bill on the streets of NYC, they believe someone will pick it up. If you agree, then you accept some things can be predicted.

WRM & JDD contend their predictions are like the \$100 bill: they forecasting the impact of incentives on behaviour.

In a book written in 1996, they predict most jobs moving online, freeing people to work from anywhere. They predict cryptocurrency:

*"This new digital form of money is destined to play a pivotal role in cybercommerce. It will consist of encrypted sequences of multi-hundred-digit prime numbers. Unique, anonymous, and verifiable, this money will accommodate the largest transactions. It will also be divisible into the tiniest fraction of value. It will be tradable at a keystroke in a multi-trillion-dollar wholesale market without borders"*

This predates the Bitcoin white paper by 13 years. William Rees-Mogg, editor of the Times\* and father of Brexiteer Jacob Rees-Mogg, was not a technologist - he [avoided word processors and wrote by hand](#) - but he and Davidson got this really, really right.

\*of London, not America's blog of record.

The thesis continues:

For centuries, you had little choice over where you could be geographically, so most people had to endure their government.

The Information Age means you can create value online.

This makes it possible to leave a jurisdiction without changing your job.

This frees you to leave a government without changing your career prospects.

This means you can shop around for governments you like.

If you are a productive individual in a high tax country, you can reduce your tax bill by moving abroad.

Many people will follow their incentives and leave.

This means western governments will face budget deficits: many wealthy citizens will cease paying tax, but it will be politically unpalatable to reduce state spending. Governments will flail around before inflating away their debt. This will have bad consequences.

Meanwhile, those who left will be doing just fine. They urge you to be among that group, which they call Sovereign Individuals.

## History

WRM & JDD assert most books about the future are about the present. They aim to buck this trend by *"making this book about the future first of all a book about the past"*, and analyse what they call 'Megapolitics' in four categories:

- Topography. E.g. no government has ever exclusively controlled the oceans.
- Climate. There was a worldwide cold spell around 1650, the same time as a worldwide period of revolutions. WRM & JDD believe this was causal.
- Microbes. The advent of sexually transmitted diseases drove prudish sexual mores. Lower death rates made battle deaths less tolerable, so better sanitation reduced governments' willingness to go to war. I'm writing in 2021, and can confirm microbes have dominated political discussion for most of the last year.
- Technology. They assert that "all men are created equal" was true in 1776: a farmer with a hunting rifle was armed as well as a British soldier. By comparison, in 1276 a peasant had no chance against a well-armed knight.

## Nation-states

WRM & JDD say the church was *the* primary institution in western Europe, then (circa 1500) declined in favour of the nation-state. They attribute this to competition and information, ie the printing press, and the subsequent rise of Protestantism (Martin Luther's works accounted for over a third of all German-language books sold between 1518 and 1525) led to competition among denominations.

For the last 500 years there was no meaningful competition among nation-states. The Internet means you can work remotely, which means you can leave, which forces governments to compete for your business. They predict nation-states will decline in significance.

Education was once managed by the church (*The Church sponsored universities and provided the minimal education that medieval society enjoyed. The Church also provided a mechanism for reproducing books and manuscripts, including almost all contemporary information about farming and husbandry*), then by the nation-state (*In the words of Eric Hobsbawm, "state education transformed people into citizens of a specific country: 'peasants into Frenchmen'"*) and they predict it is soon to be managed by technology.

They predict people will educate themselves not in what they can do for their country but rather what one horrible human in the outgroup did today: *As individuals themselves begin to serve as their own news editors, selecting what topics and news stories are of interest, it is far less likely that they will choose to indoctrinate themselves in the*

*urgencies of sacrifice for the nation-state.*

People's sense of kinship with fellow-passport-holders has declined over time in favour of other in-groups [*"A London investment banker will probably feel more at home in Seoul than he will in Glasgow"*]. Attacks on shared cultural symbols are likely a mistake for politicians who favour more redistribution. If you succeed in making people feel negatively about their country, they probably feel less good paying tax to the government. The authors consider, but think unlikely, the hypothesis that grievance studies is deliberately designed to do this:

*"If Lasch's argument is to be believed, the purpose of heightening a sense of victimization was to undermine nations, making it easier for the new, footloose information elite to escape the commitments and duties of citizenship. We are not entirely convinced that the new elite, especially most of those in the mass media, are cunning enough to reason to such a posture. It would almost be reassuring to feel that they were."*

They view nation-states with a mega-metropolis (England, France) as likely to last longer than those with many big cities (Canada). They cite Adam Smith: *Were the streets of London to be lighted and paved at the expense of the [national] treasury, is there any probability that they would be so well lighted and paved as they are at present, or even at so small an expense.*

Singapore-style city-states, rather than nation-states.

## Government as a service

Voters who disliked their government can choose between voice (writing to your mayor) and exit (moving to another jurisdiction). They say voice is becoming harder, and quote Milton Friedman: *Parents could express their views about schools directly, by withdrawing their children from one school and sending them to another, to a much greater extent than is now possible. In general they can now take this step only by changing their place of residence.*

Exit is becoming easier, and this means high-tax regimes will run into trouble: *Leading nationstates, with their predatory, redistributive tax regimes and heavy-handed regulations, will no longer be jurisdictions of choice. Seen dispassionately, they offer poor-quality protection and diminished economic opportunity at monopoly prices.*

WRM & JDD view a long term trend towards viewing government as a provider of services, and a consequent paradigm shift to shopping around for governments you prefer.

*Suppose the phone company sent a bill for \$50,000 for a call to London, just because you happened to conclude a deal worth \$125,000 during a conversation. Neither you nor any other customer in his right mind would pay it. But that is exactly the basis upon which income taxes are assessed in every democratic welfare state.*

[The top 1% of taxpayers in the US pay 38% of federal income tax. In the UK, 5% of people pay more than half of all income tax.](#) Not many of that group need to move

before governments have real trouble financing their spending. I expect some high earners will continue paying taxes - they like their country, or like being near family, or just don't mind subsidising the Queen's corgi collection. But others...won't.

The US taxes citizens abroad and doxxes anyone renouncing their passports. It had ~500 citizenship renunciations/yr in the 90s, had risen to ~4000/yr pre-covid, and [hit an all-time high in 2020, the year remote working took off](#).

Not all government-choosing is tax-based; they comment that *“those who like clean streets and resent finding gum under tabletops will find Singapore fetching. Fans of Beavis and Butthead won't.”*

I view this as a good thing! Maybe psychedelics users move to Amsterdam, and people who hate alcohol move to Saudi Arabia and everyone is better off.

## Inequality

WRM & JDD observe that some people are forced to subsidise others. They reference a model which:

*“divides the entire population of the nationstate into two classes: taxpayers, who contribute more to the cost of government services than they consume; and tax consumers, who receive benefits from government in excess of their contribution to the cost”*

They have a unique take on Marxism: they think capitalists of the world will unite and throw off their chains!

*Far from depending upon the state to discipline the workers, as the Marxists imagined, the ablest, wealthiest persons were net losers from the actions of the nation-state. It is clearly they who have the most to gain by transcending nationalism as markets triumph over compulsion*

[I'm reminded here of a Victor Mollo quote *“It's high time that someone spoke up for the strong against the weak. The underdogs have had it all their own way far too long.”*]

WRM & JDD acknowledge this will suck for people who “do not excel in problem-solving or possess globally marketable skills”. They predict that group, “leftbehinds”, will turn to nationalism and nostalgia, and “seek to thwart the movement of capital and people across borders”.

*From Slobodan Milosevic in Serbia to Pat Buchanan in the United States to Winston Peters in New Zealand, to Necmettin Erbakan of Turkey's fundamentalist Islamic Welfare Party, demagogues will rail against the globalization of markets, immigration, and freedom of investment.*

This makes sense to me. Many people do not have marketable skills. They dislike this. Their votes reflect their dislike. Some politicians will pander to those voters.

I'm not sure what politicians **can** do. “Increasing taxes and making the wealthy pay more” doesn't solve the problem that wealthy people can opt out of paying tax.

Multi-national agreements to tax people could in theory work if every single country -

Dubai, Singapore, the Caymans, etc - agrees to implement a minimum income tax.

Even assuming no tax loopholes, no special crown dependencies, no increase in anonymous cryptocurrency transactions etc...that seems unlikely. Switzerland famously ignores multi-national institutions. Its highest income tax rate bracket is 11% (kicking in above \$1,000,000 annual income) and it offers the mega-wealthy the chance to negotiate a flat tax.

## The next steps

0) the central thesis: many productive people will leave and this will cause western welfare states to become unsustainable.

1) Western governments will try to prevent this. *For the same reasons that the late, departed Soviet Union tried in vain to suppress access to personal computers and Xerox machines, western governments will seek to suppress the cybereconomy by totalitarian means*

2) They will fail. *"It will probably take a slow, painful tutorial in the realities of the cybereconomy before OECD populations are weaned away from expectations of being able to compel income redistribution on a large scale".*

[Possibly my favourite metaphor ever] *"The Don Quixote of the twenty-first century will not be a knight-errant struggling to revive the glories of feudalism, but a bureaucrat in a brown suit, a tax collector yearning for a citizen to audit."*

3) This may be unpleasant for wealthy citizens seeking to leave. *The twilight of state systems in the past has seldom been a polite, orderly process. But they will manage: The ineffectiveness of efforts to bar illegal immigrants convincingly shows that nationstates will be unable to seal their borders to prevent successful people from escaping. The rich will be at least as enterprising in getting out as would-be taxi drivers and waiters are at getting in.*

4) Western governments will have huge unfunded liabilities and a shrinking tax base. Effects will include inflation and a rise in real interest rates.

After reading TSI, I think the welfare state will break down, the result is undesirable, and I don't know what can be done about it. Maybe private philanthropy will fund the welfare state - there's a trend for billionaires to give more to charity. I hope we can spend less time bashing wealthy philanthropists - [if philanthropy earns denunciations in the opinion section of the Guardian...we'll probably see less of it.](#)

Silver lining: they seem confident there will be fewer resources devoted to state violence.

## Random passages I liked

*In the 1980s, it was illegal in the United States to send a fax message. The U.S. Post Office considered faxes to be first-class mail, over which the U.S. Post Office claimed an ancient monopoly. An edict to that effect was issued reiterating the requirement that all fax transmissions be routed to the nearest post office for delivery with regular mail. Billions of fax messages later, it is unclear whether anyone ever complied with that law.*

[I would read the hell out of a blog post on this]

*Ironically, the surge in concern about "discrimination" coincided with the early stages of a technological revolution that is bound to make actual arbitrary discrimination far less of a problem than it has ever been before. No one on the Internet knows or cares whether the author of a new software program is black, white, male, female, homosexual, or a vegetarian dwarf.*

[fact-check: Bitcoin is worth a trillion dollars and we don't know if Satoshi was a vegetarian dwarf]

*Karl Marx believed in the struggle for survival just as much as Charles Darwin, but he believed it was a war between social classes, themselves formed by economic forces. Adolf Hitler believed in the struggle for survival...But he believed that the struggle was one between different races. Marx, Lenin, Stalin, Mao, and Hitler can all be called social Darwinists, in that they saw the struggle for survival, "Mein Kampf" as Hitler called it, as the central political issue. The Marxists saw social classes as though they were separate species; the Nazis saw races in the same light*

I...have heard Communists compared to Nazis, but this justification is novel to me.

*Private competing currencies circulated in Scotland from early in the eighteenth century until 1844. During that period, Scotland had no central bank. There were few regulations or restrictions on entry into the banking business. Private banks took deposits and issued their own private currencies backed by gold bullion.*

[they may take our lives, but they cannae take our freedom to issue private currencies! 1]

*Rome was already so degenerate by the later decades of the fifth century that its "fall" genuinely eluded the notice of most people who lived through it....Many more decades passed, perhaps centuries, before there was a common acknowledgment that the Roman Empire in the West no longer existed. Certainly Charlemagne believed that he was a legitimate Roman emperor in the year 800.*

[Tired: Trump is secretly in office and planning his return.

Wired: the Holy Roman Empire has been secretly continuing for *three hundred years*]

## Other predictions

First, stuff they got wrong.

WRM & JDD expect commerce online to be done in English. They view a shared language as important to nationality, so they expect more secessionist movements, driven by language & economics - specifically Belgium and Canada (divided by language) and Italy (the North subsidises the south). That prediction looks mediocre! I write from Britain, which left a larger bloc recently, and has developed an active Scottish separatist movement. But neither was driven primarily by linguistic barriers (Scots understand English, and can almost speak it) or economics, so I think this has

failed. [Alistair Campbell disagrees with me.](#)

They understate the rise of China. *"The fastest-growing and most important new economy of the next century will not be China but the cybereconomy."*  
Perhaps it's too soon to mark them harshly, but this looks less true than it did in 1996.

The authors observe *"In the United States...nativist sentiment has historically been tinged with more than a slight tincture of racism."* and despite this, they believe black Americans, as "major beneficiaries of income transfers", will join blue collar white Americans as supporters of nationalism.

I thought that sounded false, but [a BBC article discussing why Trump did so well with black voters in 2020](#) observed that *85% of black Americans favour reducing legal immigration, more than any other demographic.*

My reading: 'nationalism' encompasses distinct concepts. One is patriotism, and another is opposition to immigration/globalisation. If they meant the second, they were spot-on: American blacks really are significantly more anti-immigration than whites.

Next, stuff they got right.

Leisure time online: *You may take a virtual visit to the Louvre...While you are wondering whether the Mona Lisa had trouble with her teeth, your computer could be downloading S. I. Hsiung's translation of The Romance of the Western Chamber. At times of your choosing, your personal communications system will read the text aloud like a bard of old...*

A rise in the number of people with no marketable skills

A trend towards winner-takes-all dynamics

Bots giving medical advice (*"Whatever questions doctors ask, the digital doctor will ask. It may determine that you drink too much wine, or not enough"*)

Cryptocurrency

Cyberwarfare

The gig economy: *The microprocessing revolution is sharply increasing the availability of information and reducing transaction costs...Lifetime employment will disappear as "jobs" increasingly become tasks or "piece work" rather than positions within an organization.*

Customised media (they say narrowcasting, we say echo chambers): *No longer will you be at the mercy of Dan Rather or the BBC for the news that reaches you. You will be able to select news compiled and edited according to your instructions.*

## What to do about it

The authors close with practical takeaways, headed by a Chinese proverb: *"Of all 36 ways to get out of trouble, the best way is - leave."*

Renounce citizenships that carry obligations. Move your money away from jurisdictions you don't trust not to confiscate it. Travel widely to familiarise yourself with other locations, to make it easier to decide to move some day. They suggest a business area: *"because incomes for the very rich will rise faster than for others in advanced*



*economies, an area of growing demand will be services and products that cater to the needs of the very rich.*” Place your business online and prepare to transact in cryptocurrency.

This piece of advice alone was insanely strong. I sorta wish I’d read this book a decade ago and invested in Bitcoin so that I could now be boring my friends by saying “If I’d held onto my Bitcoin from 2011, I’d be a gazillionaire”. Instead I bored my friends by telling them how much this book predicted. One thought cryptocurrency wasn’t that far off in the late 90s and pointed out it was the original goal of PayPal.

My response: The Sovereign Individual was published Feb 1st, 1997. [Peter Thiel said it was the book he is most influenced by](#). Paypal was founded in December 1998. The name of Paypal’s CEO? ~~Albert Einstein~~. Peter Thiel.

They say the surest way to predict the future is to create it. Perhaps the second surest is to alert Peter Thiel to your predictions and let nature take its course.

# The Sovereign State and Its Competitors by Hendrik Spruyt (1994)

I really like how clear this book's writing is, even though it's designed for a scholarly audience. I think that the best way to start this book review is with the first paragraph from the book's introduction:

"History has covered its tracks well. We often take the present system of sovereign states for granted and believe that its development was inevitable. But it was not. The sovereign, territorial state had its own peculiar rivals that very well might have held the day. Now that dramatic changes within and between states are taking place, it is appropriate to rethink our explanations of the origins of the state system, analyze the forces that shaped it, and reflect on the possibility of its demise." [1]

Our current international system consists of sovereign states.

A sovereign state is an institution which claims complete authority within certain lines on a map and no authority outside of those lines. It has a bureaucratic hierarchy within its borders and recognizes other states as equals outside its borders. This idea is so ubiquitous that we don't stop to think about how weird of a claim that is or what other kinds of international systems could exist.

Sovereign states developed in Europe out of feudalism during the High Middle Ages. During this time, there were other competing systems that vied to replace feudalism: city-states and city leagues. These different systems competed with feudalism and with each other until the system of sovereign states became dominant. This process was finalized with the Peace of Westphalia (1648).

The Sovereign State and Its Competitors focuses on three regions of Europe: France, Germany, and Italy. They exemplify the development of the sovereign state (in France), city leagues (in Germany), and city-states (in Italy), between the years 1000-1400. My main criticism of the book is that it is too short. I would like to also hear Spruyt's analysis of the Swiss cantons, which might be most similar to a city league, and the English parliamentary sovereign state.

---

The first chapter is a literature review (this is a scholarly book). The second chapter describes Spruyt's theory of system change in the context of other theories of system change. If you are a casual reader, you can skip both of these chapters and start with Ch. 3.

Before we look at how the sovereign state began, we should look at the institutions which existed in Europe beforehand. This is now referred to as feudalism, although that term wasn't invented until later, and it suggests that there was a single system, instead of multiple: within the nobility, within the church, among the peasants, within the cities, and between all of these groups.

Politics among the nobility was based on personal relationships, formalized in oaths of fealty. A lord would have multiple vassals, and the obligations in each relationship were specified in that oath. Each vassal could have their own vassals. These relationships were non-exclusive and only weakly hierarchical. An oath of fealty to one lord did not preclude an oath of fealty to another - although you might have problems if those two lords went to war. A single individual could hold multiple titles, at different levels of the hierarchy, which implied different vassalage. Henry II, as duke of Normandy, was a vassal of the French king, but the same Henry II, as king of the English, was not. The right to rule was based on personal oaths, not a broader notion of sovereignty. These oaths were rarely directly broken (although people certainly did interpret them creatively) because no one would follow an oathbreaker. Since oaths were inherently religious, excommunication would also revoke all oaths involving that lord.

The church had its own hierarchy, modeled on the administration of the Roman Empire. Archbishops and bishops were the spiritual leaders of the provinces and dioceses of the Roman Empire. Bishops and monasteries anchored the urban centers that survived the fall of the Western Roman Empire and preserved much of the ancient learning. The church claimed a divine right to rule over all Christians. The sovereignty of the church was based on the sovereignty of God. Christianity should be taken to all people, so this is a universal claim. However, Christ's kingdom is not of this world, so it is not clear how this claim pertains to political power.

Most people were part of the third estate, peasants, rather than the first two estates, warrior/nobility and clergy. Peasants were typically in subservient "manorial" relationships, where they had to pay regular taxes/rents and were legally inferior. While both the nobility and bishop could enter into feudal relationships, the peasants were forbidden from taking oaths of fealty or bearing certain weapons. While manorialism looks like local sovereignty, since serfs were legally bound to the land, it can be thought of either as territorial rule or personal rule.

On Christmas Day in 800, the Pope crowned Charlemagne emperor, promising the revival of a theocratic Roman Empire. This title was later inherited by the Germans. The Holy Roman Emperor claimed the political authority over Christendom, much like the Pope's universal claim of religious authority. The German Emperor regularly tried to demonstrate his supremacy over Italy and only occasionally over France & England (most notably with Richard the Lionheart). This belief system was common in ancient empires. While they were aware that there were practical limits to their power, in principle, they claimed a divine right to rule over all people. Ancient empires typically would rule over their entire economic sphere - both because that was a major military goal of the empire and because they made trade within the empire much easier

than trade outside of the empire.

The distinction between religion and politics was much less clear than it is today. Bishops had significant economic and political power and the emperor claimed theocratic rule. Having two universal claims was bound to lead to problems. These problems came to a head with the Investiture Controversy, which started in 1076, over whether the Pope or the Emperor had the right to appoint bishops. The controversy lasted for decades until a compromise was reached (the Pope appointed bishops, with the Emperor's approval). These disputes helped to distinguish what issues were considered religious and what issues were considered political. The Pope was able to gradually solidify a church hierarchy centered on Rome. Politically, there was more fragmentation in both Italy and Germany, since local powers could use the competition between the Pope and Emperor to their advantage.

---

Although most historians refrain from using the term "Dark Ages", if you do, the Dark Ages were from about 500-1000 A.D., excluding Charlemagne's reign. The claim that the Dark Ages lasted until the Renaissance is myth written by men during the Renaissance who wanted to feel superior to those who came immediately before. The High Middle Ages (1000-1250) marked the end of the Dark Ages:

- The population of Europe increased dramatically.
- Existing cities grew and hundreds of new cities were founded, especially in Northern Europe.
- Trade expanded: luxury goods in the Mediterranean and bulk goods in the North & Baltic Seas.
- The economy became more monetized.
- Grand Gothic architecture was built, including the Notre-Dame Cathedral of Paris.
- Learning was institutionalized in the first universities.
- Major intellectuals of this era included Thomas Aquinas and Fibonacci.

The new towns not only differed from the rural feudal system institutionally, they also differed in worldview.

This can be most dramatically seen in their understanding of time. Time was largely irrelevant to most rural activities. Only the church measured time, to determine when religious services and holy days occurred. Urban businesses used a different idea of time, one precisely measured by mechanical clocks, an invention of man. These two standards of time were literally raised over most medieval towns. The two tallest buildings were the church's bell tower and the secular clock tower.

Understanding time differently leads to more material conflicts. Usury was forbidden by the church because it was understood as charging people for time, which is a free gift from God. A similar ban was placed on paid teaching because knowledge is also a free gift of God.

There were contradictions in many other ideas as well:

- Depersonalized business relationships vs personal oaths of fealty.
- Monetary transactions vs reciprocal gift giving.
- Predictable written law (often based on Roman law) vs local customs and traditional law.
- Trial by evidence vs trial by combat.
- Status based on wealth vs status based on ancestry.

Most of the towns during the High Middle Ages were republics that mixed democratic and aristocratic institutions. The town government was typically elected by the burghers (citizens of the city). The qualifications to become a burgher varied - sometimes property ownership was required and sometimes you just had to live in the town for a year and a day. Towns were places of relative freedom, especially compared to the feudal obligations of the rural peasants.

The new international systems of the High Middle Ages developed as a consequence of the interaction between the previous feudal systems and the increasingly powerful towns.

---

## France

The sovereign state developed in France from an alliance between the towns and the king.

French towns did not grow as large as Italian or even German towns during the High Middle Ages. While they were not large enough to be politically independent, they were significant, especially economically. The French king promoted the political interests of the towns in exchange for increasing monetary taxes, which allowed him to further consolidate his power.

While businesses like lower taxes, it is even more important for the taxes to be predictable. The feudal tax system was especially bad for business. Because feudalism has non-exclusive rule, a town could be obligated to pay taxes to multiple lords, who collected the tax when they saw fit. Towns would much rather pay taxes directly to the king once a year, although the amount to be paid remained under continual negotiation.

The king and towns also agreed on promoting a more consistent legal regime based on Roman Law and on standardizing weights and measures, although these efforts were not completed until much later. When there were conflicts between a town and a local lord, the king would often intervene on the town's behalf, and bring the town into territory directly controlled by the king.

Feudal obligations to the king were typically in food or in military service, offered in exchange for inheritable land. The revenue the king got from the towns was money. Money has a lot more uses.

The main thing that the French king used money for was mercenaries. He was no longer reliant on the knights of potentially disloyal feudal lords. Lords were allowed to choose to contribute money for mercenaries instead of appearing in battle themselves.

The monetary taxes from towns also meant that the French king was less reliant on feudal revenue sources. The king would give the nobles tax exemptions or even pensions in return for compliance and land. Once this system was established, the French king did not have problems with the nobles conspiring against him because they had been bought off.

The king could also establish a bureaucracy, paid in money, who reported directly to him. At the lower levels of the bureaucracy were tax farmers, local people who got a certain percentage of the revenue they raised. This aligns their incentives with the goals of the king. The higher levels of the bureaucracy were appointed and paid directly by the king. They were always from a different part of France from where they worked and moved regularly so they wouldn't follow local interests. Most of the members of the bureaucracy were from the towns (or lower nobility) because they were literate and often had administrative experience. This further cemented the alliance between the king and the town and made the old feudal hierarchy irrelevant.

The church was initially an ally of the French King because they were common enemies of the Holy Roman Emperor. The church also provided a model of centralized administration that the king could borrow from. However, the logic of a sovereign king and of a sovereign church are contradictory and led to conflict. By the time this conflict occurred, the French king had already cemented his power and was able to force the Pope to recognize his right to tax and even arrest bishops and monasteries.

By 1400, the French king had consolidated a sovereign, territorial state.

## Germany

Germany had a weakening emperor, powerful lords, and growing leagues of cities.

German towns were in many ways similar to French towns. They were similar in size to the French towns. But there were a lot more of them. German political and religious leaders founded many new towns, especially as the Germans spread eastward, using groups of people from already established towns. Both French and German towns mostly traded bulk goods like wheat, fish, lumber, furs, and wool. But German towns traded a larger volume of goods for longer distances across the Baltic and North Seas. Both French and German towns wanted

predicable taxes, consistent legal systems, and standardized weights and measures. But only the French king provided them.

The German emperor was ideologically committed to universal rule of the Holy Roman Empire. To accept that his rule should be limited only to Germany would require significantly limiting his worldview. To maintain this idea, the German emperor spent much of his time in Italy, trying to keep the Italian towns under his control. Any success he had there was opposed by stronger cooperation among the usually fractious Italian cities and with the Pope.

In order to campaign in Italy, the emperor relied on the support of the nobility. The nobility were able to claim increasing control within their own territories, including minting coins with their own facsimile, writing and enforcing laws without appeal to the emperor, setting their own weights & measures, and controlling the local towns. By 1356, the emperor officially agreed that the nobility had the right to elect his successor (instead of just approving his choice), which ensured their long-term dominance.

German towns did not like being subservient to feudal lords any more than French towns did. Bishops were particularly disliked. Not only did the towns disagree with the bishops on political and economic issues, they also disagreed on moral issues. The church forbade many business practices and found all forms of monetary transactions suspect. People should not be putting that much effort into gaining wealth.

Towns would sometimes look to the emperor for support, and occasionally he would grant it, but mostly, the towns were on their own. Individual towns were not large enough to fight the increasingly powerful lords.

So they formed leagues.

When a town wished to gain more independence, its residents would (illegally) swear oaths of mutual assistance, elect a town council, and start trying to negotiate or appropriate rights and liberties from their lord. A city league formed in a similar way. Representatives from groups of towns would gather together to coordinate defense. Once established, the league would reduce tolls, regulate weights, measures, and coinage, standardize justice, and negotiate or wage wars with the lords.

By 1358, the various German city leagues had merged into the Hanseatic League. The Hansa started as a traders guild that would meet on the island of Gotland in the Baltic Sea. At some point, the towns started sending representatives instead of individual traders.

At its height, the Hanseatic League contained about 200 towns - about 77 as full members and the rest as associates. The number is uncertain because the various towns of the Hansa had different official status. Some were free and imperial cities, which recognized no sovereign but the (uninvolved) emperor. Some were officially under the rule of a lord, but in practice set their own policies. Some local lords had a closer alliance with their towns. There were also four

foreign cities (London, Bruges, Bergen, and Novgorod) which had a permanent Hanseatic trading post. Like an empire, the League sought to dominate its entire economic sphere, although the League spread by inviting new members, rather than by conquest.

The main institution of the League was the Hansetag, a meeting where all of the towns sent delegates to determine policy for the League which occurred every few years. Each town had one vote, but in practice, the major towns, especially Lübeck, dominated the League. While towns could delegate other towns to represent them, decisions made by the Hansetag were binding even for the towns that did not attend. Refusing to obey the decisions of the Hansetag would result in fines or even expulsion from the League. Along with the Hansetag, there were three or four regional groups for more local issues.

The state capacity of the Hanseatic League was similar to the contemporary sovereign, territorial states.

The Hanseatic League could raise revenue by requesting fees from the various cities. It also set tolls (and forbade others from doing so) for ships entering and leaving harbors.

The Hanseatic League monopolized trade in the North and Baltic Seas. Only Hanseatic ships could enter Hanseatic ports. They had extensive trade and tax privileges in the countries where they had permanent foreign offices. The heads of these foreign offices were authorized to negotiate treaties on behalf of the League.

The Hanseatic League could and did wage war on the same scale as sovereign states. After winning a war against Denmark, they controlled succession of the Danish king and the castles along the shipping routes. Blockade was a more common form of coercion than war. It was easier to force concessions from kings by refusing to sell food in Norway or boycotting Flanders than by going to war.

The Hanseatic League made the laws more consistent between cities. Some of this consistency occurred because new cities were often formed by groups from old cities, who brought their laws with them. Lübeck Law later spread across much of the League. Even in towns with variations in local law, Hanseatic merchants could expect to be tried according to the traditions of the League.

State capacity can also be seen from uniform culture. States don't just standardize weight, measures, and coinage, they also standardize language and customs. French only became a single language, distinct from Spanish and Italian, through the consistent efforts of the French state. During the 1300s and 1400s, the Hansa's dialect of Low German spread across a larger territory than the Standard French of Paris. Moving between towns was easy for Hanseatic merchants, but marrying outside the league was forbidden.

In 1400, the German city league was as successful of an institution as the sovereign, territorial state.



# Italy

Italian cities were strong enough to act as independent city-states.

The Italian cities were much larger than either the German or French towns. Only Paris came close to their size. Individual cities obtained more revenue and fielded larger armies than burgeoning states like England or Portugal.

Italy had remained more urbanized after the fall of the Western Roman Empire and retained more of Rome's social structures and law. Additionally, Italy's urbanization was due to differences in trade.

Trade in the Mediterranean had a different nature than trade in the North and Baltic Seas. Merchants made most of their income from luxury goods, like spices and silk. Luxury goods had significantly higher profit margins, which allowed cities to gather more wealth, and people. Some bulk goods, especially grain, were also transported long distances. The different emphasis can be seen in the different types of ships: Mediterranean ships had less cargo space, but a larger crew, including professional soldiers. Most of a city's wealth for the year came from a few ships traveling to a few cities in the Eastern Mediterranean. Much of the conflict between and within the Italian cities can be traced to trying to secure the exclusive rights to trade for a small volume of luxury goods.

The distinction between rural aristocracy and urban burghers did not exist in Italy because the aristocracy was also urban. Early on, the urban social classes allied to throw off imperial or ecclesiastical rule and established city-states. Instead of a bishop ruling the city, the aristocracy would make deals to choose the bishop - or even the Pope. Although the independent Italian cities mostly remained republics through the rest of the Middle Ages and Early Modern Period, they tended to become increasingly aristocratic.

Factions were extremely important in Italian cities. Extended aristocratic families, guilds, and neighborhoods could all have their own armed forces. Government existed for the benefit of the currently ruling faction, not for the city as a whole.

Foreign interference was also more important. The Byzantines, Arabs, Germans, Normans, French, and Spaniards all ruled part of Italy at some point during this time period. Southern Italy was almost always under foreign control, remained feudal, and didn't develop large or independent cities.

The city-states gradually consolidated as larger cities forcefully subjugated their neighbors. They did not, however, integrate them. Subjugated cities retained their own local customs and

laws, but not independent military or trade. Rural areas had little or no political influence.

The combination of all of these factors makes Italian history extremely complicated. Even with all of the institutional diversity of the city-states, they did not develop the logic of the sovereign, territorial state. The cities did consider themselves sovereign and did recognize limits to their territorial control. They did not homogenize institutions or culture within their territory. The government did not try to rise above the fray and represent the interest of the entire people instead of a particular faction. They never established a monopoly on violence or justice.

In 1400, the Italian city-states were as powerful as the strongest sovereign states in Europe.

---

Over the next 250 years (1400-1650), the sovereign, territorial state system came to dominate Europe.

How did the state system come to dominate over city leagues and city-states?

One of the most common explanations is war. Sovereign states were more effective at waging war and eliminated rival systems through conquest. There are several problems with this explanation:

- While sovereign states were militarily superior to feudal systems, they weren't superior to either city leagues or city-states. When these institutional alternatives were at their height, they successfully waged war, even against the largest states.
- The sovereign state system did not eliminate small actors, as long as they conformed to its logic. Individual Hanseatic cities remained independent, long after the League stopped meeting. It was not until the French Revolution and Napoleon that an army marched through Europe and forced everyone to adopt similar institutions.
- Even if war were a good immediate explanation, the question of why states were more effective at waging war remains unsolved.

One reason for the sovereign state's success was that it proved more effective at standardizing weights and measures, coinage, and judicial systems. Almost every town and major lord in Europe had their own measurement system and currency, which dramatically increased transaction and information costs for merchants. Moreover, there were a variety of legal systems, some codified and some entirely based on traditions, so there was uncertainty in how to redress wrongs. The French king continually encouraged people to exclusively use his standards, eliminating local variation. The confederated cities of the Hansa and the contested sovereignty of the Italy city-states were less effective at standardizing their weights and measures, currency, and laws. I am not as convinced by this argument as Spruyt is, both because the Hanseatic League also continually tried to create standards (Lübeck or Magdeburg law, Lübeck marks, and Rostock barrels) and because the French did not complete this process

until the French Revolution. If the Hanseatic League had persisted, they might have standardized to a similar extent by 1800.

Another more important reason for the decline of the Hanseatic League is that it was unable to be as effective in international diplomacy. While the League did sign treaties, they were enforced by the individual city councils. The League's tools to coerce the individual cities (fines, threats of expulsion) were limited. Individual cities would try to take advantage of the benefits of a treaty without respecting the limitations or would try to negotiate more preferential treaties with other powers on their own. This undermined the cohesion of the League. To try to avoid other powers' attempts to divide them, the League refused to establish a list of what cities were members. This further undermined the League's negotiating power. A king with final authority over his subjects could convincingly bind his entire society to treaties, allowing for more complicated international relations to develop.

Some of the League's diplomatic problems were due to how it was organized. But some of the problems were because sovereign states preferentially dealt with other sovereign states. They would purposely empower other actors that behaved like sovereign states. This encouraged the elites in other societies to begin mimicking the sovereign state system. If they claimed regalian rights, other states would recognize them as sovereign and help them to enforce these rights. The Italian city-states were adopted into the sovereign state system as small sovereign states. In Germany, this process culminated with the Peace of Westphalia which ended the Thirty Years War. Established states like France and Sweden recognized the sovereignty of the German lords like Saxony and Bavaria. They could even decide questions of religion within their realms. The Holy Roman Emperor became a figurehead. Most of the Hanseatic cities were forcibly incorporated into these new states. A few of the largest became small states themselves. The

Hansetag met one more time, but only three cities sent representatives, and they decided to disband. Two of these cities (Hamburg and Bremen) are still considered Free and Hanseatic Cities within Germany's federal republic - Lübeck's status was permanently revoked during WWII. Once the international system of sovereign states became partially established, it promoted itself over other rival systems, until it eventually spread over the entire world.

---

Spruyt concludes by restating his general theory of change in international systems.

Most of the time, the international system is stable. Conflict occurs between and among actors, but what kind of actors exist doesn't change. For example, both the Napoleonic Wars and World Wars changed the relative status of different states, but they did not change the fact that international politics is dominated by sovereign states. Change in the international system only occurs occasionally, so its evolution can be thought of as a punctuated equilibrium.

Occasionally, something changes that makes the current international system untenable. The increase in trade in Europe starting in 1000 is one example.

This change occurs in two steps.

First, political entrepreneurs create new institutions to meet their needs. This process is non-deterministic: different regions tend to develop differently. What alliances form between different interest groups determine what innovations occur. In Europe in the High Middle Ages, this stage of systems change led to the formation of the sovereign state in France, city-states in Italy, and city leagues in Germany.

Then, there is competition between the different sorts of institutions that have arisen. Although each of the new institutions is more effective than the previous system (feudalism), it is not clear which will become dominant over the others. This second stage of systems change occurred in Europe from 1400-1650 and ended with the dominance of the sovereign state system.

Spruyt thinks that we might experience another change in our international system soon. The sovereign state isn't as dominant as it has been in the past. A regional association (the European Union) offers one alternative. Multinational companies and finance mean that international economic policy is no longer governed by relationships between states. Religious extremist groups explicitly reject the state system. Many of the threats we face do not respect borders. Although this was written in 1994, his examples are even more salient now. These trends do not mean that the international system of states will fail, but it does suggest the possibility. By studying how these changes occurred in the past, we can be more prepared if the international system starts to change in our lifetime.

# The Suicide of Reason by Lee Harris

## I. The Alien Space Ship

It seemed like such a nice day when you and me and Scott and the rest of the SSC readership were kidnapped by alien slavers. En route to its destination, the space ship suffers a catastrophic fault and crashes, stranding us on an alien and hostile world. Some bright spark points out that we should unite for mutual protection. We agree that this is in our best interests, at least right now. But what if something fanged and hungry leaps onto Scott? Suddenly, it's in his best interest that we all defend him, but it's in our best interest to run like hell. We know this, so how can we have an agreement that's more than mere words?

It's worse than that. We don't need to worry about external predators, we can be just as dangerous to each other. Right now, when we're all equally broke, we can agree to respect each other's rights, but when one of us manage to catch some food and the rest of us don't, we have a rational incentive to steal it from him. How do we stop ourselves going all Lord of the Flies on each other?

Fortunately, something else survived the crash: a crate full of collars the slavers use to keep their cargo in line. The collars can be programmed to inflict pain on their wearers when they do some things, and make them feel good when they do others. What it is that makes the wearer feel good or bad is up to the programmer.

So here's our solution. We seal our compact by putting on the collars and programming the collars to make us feel absolutely wretched if we do certain things, like running away when there's danger to the group. Then we break the keys and smash the programming terminals so no one can take their collar off or quietly reprogram them. We don't need to rely on people's words; we know that their actions are constrained by the collars. The chance of us sticking together in the face of danger just went way up.

This is the core idea of Lee Harris extremely interesting and incredibly disturbing book, *The Suicide of Reason*. He argues that we have indeed been dropped on a hostile planet by an alien spaceship and left with programmable shock-collars. The hostile planet is called Earth, the space ship is called Evolution, and the collars are called Shame. Social cooperating is founded on shame; we don't stick by our social compacts because we've carefully thought it through, decided it's right and committed to sticking to it, we stick to it because we're inhibited by the pain that shame inflicts if we don't.

"In a strong group, when an individual is given a chance to desert his fellows in order to save his own skin, he will be inhibited from this act of selfish betrayal by an unbearable visceral shame.

What will keep him loyal to the group is not his higher faculties of reflection and cognition - all of which may be screaming to him, 'Run for your life, you fool!' Rather, it is the physiological reactions that have been programmed into him from an early age through the process of shaming."

Shame is powerful - so powerful that people kill themselves rather than endure the pain shame inflicts. According to Psychology Today, 30,000 Japanese people kill themselves yearly, mainly out of shame.<sup>1</sup> A quick search for "shame + suicide" finds countless more examples. This makes sense; if the point of shame is to hold control our behaviour, even in the face of death, then it can only work if it is something we naturally fear more than death.

Harris credits this idea to the essay Evolution and Ethics, by T.H. Huxley, "Darwin's Bulldog". Huxley begins by describing what he calls the Cosmic Process. The Cosmic Process is the indifferent laws of nature going about their business. The process isn't cruel or evil; it just doesn't care. If a baby falls to his death, that isn't because the baby deserved to die, it's just how gravity works.

The part of the Cosmic Process that particularly concerns Huxley and Harris is Darwin's struggle for existence. The same indifference applies - if a badger eats a litter of cute kittens, it's not because they deserve to be eaten, it's just the way things are. Losers die, winners survive a little longer, and that's all there's to it.

For most of our species' history, human beings have been just as much subject to the Cosmic Process as anyone else. Tribes wipe each other out, within a tribe stronger members prey on the weaker, all sorts of horrible things are done in the name of survival, and that's all there is to it. This leaves us with a legacy.

"[W]ith all their enormous differences in natural endowment, men agree on one thing, and that is their desire ... to do nothing but that which pleases them to do, without the least reference to the welfare of the society in which they are born. That is their inheritance (the reality at the bottom of the doctrine of original sin) from the long series of ancestors, human and semi-human and brutal, in whom the strength of this innate tendency to self-assertion was the condition of victory in the struggle for existence."

Harris comments:

"Today, many quite intelligent men believe that the doctrine of original sin is sheer nonsense. But what arguments could any modern skeptic use against Huxley's version of original sin, which, unlike Augustine's, does not require us to believe in a fable about talking serpents and forbidden fruit, but simply the matter-of-fact acceptance of the law of natural selection?"

Our original sin is that we're descended from a long, long line of people who managed to reproduce because they were tougher, stronger, smarter, or just plain nastier than their neighbours, and that has consequences for our behaviour.

Yet something strange has happened. Human beings, uniquely, have turned against the Cosmic Process and opposed it. In the same way that a gardener tames a stretch of ground and encourages the plants he likes and gets rid of the ones he doesn't, we have tamed our own savage natures, encouraging those qualities we like and constraining those we don't, a process Huxley calls the Ethical Process, and we've done so with a tool the Cosmic Process gave us. Evolution didn't give us shame because it wanted us to be civilized; it did so because groups that stick together are more likely to pass on their genes and memes than those that don't. The Better Angels of Our Nature<sup>2</sup> are the Angels of Shame. We are not better people than our ancestors because we have rationally decided to be so, we are better people because we've been brought up to feel ashamed not to be so, and this training runs so deep that some kinds of behaviour are almost literally unthinkable.

At the next SSC meetup, what chance that two people accidentally insult each other and fight to the death? Or that an unaccompanied woman is raped and murdered because she's not under anyone's protection? Yet both behaviours are common throughout most of human history. In some parts of the world, they still are. That you are shocked and disturbed to even consider such questions is your shame at work. Thank it for its service.

Shame isn't just irrational, it only works because it's irrational, because reason implies choice - you choose whether or not to accept a proposition. How reassuring would it be for women for an SSC meetup to advertise with the slogan, "Don't worry! After thoroughly Steelmanning the proposition and carefully weighing the evidence, we've decided that raping and murdering you is off the table!". The whole point of shame is to deprive you of that choice, to make some things seem so obvious and natural there's no point of even thinking about them.

Since it's irrational, shame doesn't need reason to inculcate it, and you can program a shaming code quite nicely without any arguments.

"The shaming code [...] does not need to be instilled in us through etiquette books or helpful advice - in fact, there is no need for words at all, since slaps and blows and angry faces are often a much more effective way of instilling the shaming code than verbal formulae"

Military bootcamps are long on training and discipline and short on explanatory lectures, and they do pretty well in changing people's behaviour. On the other hand, how many people have had their lives completely changed by reading a self-help book (I mean really?). That looks a lot like shame: 2, reason: 0. Or to put it another way, if information was enough, we'd all be billionaires with six pack abs.

Since shame doesn't need arguments, it's irrational in what it reacts to. It doesn't have to make sense. There doesn't need to be any logical connection between the different things that trigger shame. You can be made to feel ashamed of human sacrifice, smoking, and eating meat on Fridays without ever bothering to ask, "What do these things have in common?"

In his previous book, *Civilization and Its enemies*, Harris made this explicit:

“Shame is utterly shameless. It is without principles. It does not begin with a set of ethical axioms from which it proceeds to deduce all the specific applications of these axioms in our daily living. On the contrary, the shame system makes up lists and inventories of forbidden conduct, like the code in Leviticus, where there often seems to be neither rhyme nor reason guiding the selection. It follows no apparent method and makes no connection: it just piles one thing on top of the other.”

Harris calls these lists of things that trigger shame the ‘shaming code’.

If Harris were just saying that this is our evolutionary inheritance that we are growing away from, there’d be little to worry about; he’d just be giving Pinker the missing mechanism, the same way Mendel gave the missing mechanism to Darwin’s theory. Yet Harris argues that it is not the case that we have all these shame-based irrational taboos on one side and progressive enlightened values based on reason on the other; it isn’t the case that the Cosmic Process relies on Shame and the Ethical Process on Reason. It’s shame all the way down.

We don’t reject racism and sexism because we’ve carefully considered the issues and come to a measured conclusion; we do it because we’ve been raised to regard such things as shameful - and by ‘we’ I mean ‘all of us’. Even people incensed about ‘political correctness’ and ‘Social Justice Warriors’ would be horrified by a shop sign saying “whites only”. Our Ethical Process has advanced to the point that we are sickened by things our ancestors wouldn’t have glanced twice at.

Basically, we are not rational actors analysing the world and plotting our course by the best conclusions we can get. We’re a bunch of monkeys wearing shock-collars whose programming we don’t really understand, and most of us don’t even know exists.

This is unsettling. The trouble is that it also explains a lot.

When Scott wrote a nice review of my Anti-Racist Q & A, he wrote:

“I think the most astounding part is that it might be one of the first things I’ve ever read to argue against racism. [...] I’ve read a lot of articles condemning racism, and accusing people of racism, and being very upset about the racism inherent in society. But this might be the first one I’ve ever read to argue against it.”

This is ridiculous - we live in a society that never shuts-up about racism. Surely there must be other people out there offering rational arguments against it? Yet, by Harris’s argument, this makes sense. The people condemning and accusing and being very upset aren’t arguing because they’re not relying on reason. They’re trying to maintain the shame-based tribal taboo against racism.



In the run-up to legalizing gay marriage, some conservatives said that this would lead to bestiality and paedophilia and so on. It's not that those arguments are offensive, but that they make no sense. How do you start by accepting gay marriage and end up hitting on the livestock? Yet the arguments do make sense from a shame-based point of view. Homosexuality is a taboo, paedophilia is a taboo, bestiality is a taboo. If you break one, how can we know you won't break the others? A taboo is a taboo. Shame doesn't have to make sense.

Why is political correctness so crazy? How can a black security guard be fired for simply telling a student not to call him the n-word?<sup>3</sup> Because a taboo is a taboo and shame doesn't have to make sense.

Why am I so upset by seeing cats and dogs in Chinese wet markets, and yet so indifferent when I see pigs and chickens in our factory farms, even as I can intellectually accept that they are the same thing? Well, shame doesn't have to make sense.

Why do deliberately silly rules help people cooperate and support sensible ones?<sup>4</sup> Shame doesn't have to make sense.

It also explains why it is so difficult to actually change your mind<sup>5</sup>. I've been reading Lee Harris for about a decade, and he's one of the hardest authors I've ever tackled. Reading him I found myself dawdling, procrastinating, focusing on non-essential side-points in his argument, and, why I finally forced myself to get to grips with some of his conclusions, feeling physically sick with the effort. That's to be expected if my shaming code has set me to regard Reason as the highest possible good.

(I might be over-egging the pudding here, but I think this also supports why people who reach Enlightenment end up having sex with anything that can't get out of their way <sup>6</sup>. If reason isn't enough, and you need some strong emotion to motivate action, then it makes sense that people who dissolve their shaming code default to the next strongest drive.)

This is a radical attack on liberalism, the kind of liberalism that we all accept unthinkingly, regardless of whether we call ourselves liberals, conservatives, libertarians or what have you. All of these rely on the premise that you can establish a society by reason and rational agreement. No matter how much people may have hated Bush, Obama or Trump, they determined to get rid of them by either free elections or due process of law - i.e. reasonable actions in a rational society - rather than putsch, open revolt, or the power struggle, the normal method throughout most of history. Yet if shame, not reason, is the basis of any sort of functional society, then this goes out the window.

We are not naturally reasonable, liberal people. We are not rational actors because we are inherently rational. We are rational actors because we are the inheritors of an Ethical Process that has instilled in us a shaming code that prevents us from acting unreasonably. The Ethical Process, at least in the modern West, has succeeded so well, and for so long, that we've

forgotten the nature of the Cosmic Process or that it even exists. We've worn the collars so long that we've forgotten that they're there, and the jungle that made them necessary.

## II. So, where did the process come from?

Objection: at least in the liberal West, we have created cultures of popular reason, cultures where all people - not just a few blessed by genes, status or quirks, but all people - are encouraged to think for themselves and decide what's best for them. Saying that we are this way because we have inherited a shaming code designed to produce rational actors just kicks the problem back a level. If we've inherited it unthinkingly from our parents, they must have inherited it unthinkingly from theirs, and it's turtles all the way down.

Harris is aware of this objection, and puts a ton of historical evidence against it. He points out that during the first Enlightenment, that of ancient Greece, it was taken as read that most people weren't fit for reason (Aristotle took this line). He notes that the original philosophes of the Enlightenment put all their hopes in 'Enlightened Despots' to usher in the age of reason with an iron fist. He points out that the great revolution of Reason, the French, ended in slaughter, Bonapartism, and condemning to death its principle prophet, the Marquis of Condorcet.

"We are all [Condorcet's] children; we all subscribe to his ideas; they live in us, and we live through them. To be a modern Westerner is to be a spiritual descendant of Condorcet and this is true no matter whether we call ourselves liberals or conservatives. Who in the leading nations of the West today is opposed to free public education for all children, both rich and poor? Who wishes to revive slavery or the slave trade? Who believes that secular education is an evil to be stamped out? Who is opposed to universal manhood suffrage or free and fair elections? Who wishes to deny women the right to vote?"

And yet he ended his life arrested by the Committee for Public Safety and was found dead in his cell. In a revolution dedicated to its name, the force of reason couldn't spare its greatest proponent.

Okay, but what about the other revolution, the American? Harris makes a very convincing case that America created just that culture of stubborn individualists that Condorcet hoped for, but it wasn't universal secular education that created it (Condorcet's great hope). Instead it was the fact that America was settled by Protestant Dissenters who took it as a literal matter of faith that everyone had the duty to read the Bible for himself and make up his own mind.

"Protestant Dissenters encouraged everyone in the community to be stubborn individualists. By dissenting from the opinion of those in established authority, they cultivated and passed on a visceral aptitude for intellectual independence. Dissent was in their blood, and the courage to dissent was part of their character."

The extent to which the Dissenters resisted any sort of authority can seem extreme even by the standards of modern libertarians. They would randomly dismiss their preachers, not because the preachers did anything wrong, but because they did not want to risk any individual becoming a leader. This resistance to authority, even voluntary authority, animated their hatred of Mormonism. They were afraid of any man who could proclaim himself a prophet and gather an army of followers.

This isn't the usual "Yay, Pioneer Spirit!" stuff. The Dissenters were lucky to find themselves in a place where this sort of stubborn individualism was possible, because Native Americans were stuck at the level of warring tribes and were dying from smallpox and other European diseases. That meant that the American settlers could rely on voluntary militias that would never have stood a chance against Cossacks, Janissaries, or the armies of countries like Sweden or France. Indeed, the reason the Protestants ended up in the New World is precisely because they fled the cultural predators of the Old.

Americans, in particular American conservatives, look down on Europeans as being more collectivist and statist. What they ignore is that, in a world where you are surrounded by powerful nations and empires, stubborn individualism is a prescription for social suicide. Americans became a people of stubborn individualists because they were lucky enough to do so, and they created a shaming code that exalted this character above all else.

(This also explains the paradox at the heart of the American experiment - how a country dedicated to liberty could tolerate racism and the oppression of its black minority for so long. All together now: Shame doesn't have to make sense!)

So the culture of liberty we enjoy today isn't the result of the force of reason, or an inevitable historical progress. It's a lucky accident. That's an unpleasant conclusion, but it also explains a lot. It explains why European libertarians, who usually admire America and its founders, have found it next to impossible to make any headway in their native countries. The cultural preconditions are missing.

Lee Harris is reviving an intellectual tradition that went out of style with the fall of the Berlin Wall, namely Hegelian Materialism. The core of this is that having good ideas isn't enough – "Mere goodness can achieve little against the power of nature." You can talk all day long about tolerance and individualism, and how great they are, but all you'll ever achieve is fantasy unless certain material preconditions are in place. If you want women to get the vote, you need those arguments in place, and also a giant world war that means you rely on their labour. If you want to end slavery, you need the arguments in place, and also the violent dissatisfaction of working people forced to compete with slave labour. If you want women's emancipation, you need the arguments in place, and also contraception, machine production and modern medicine. If you want tolerance, you need the arguments in place, and also the practical demonstration of why Intolerance Is Bad in the form of religious wars that killed half a continent. You forget this at your peril. Writing, arguing and hoping for reason and tolerance and progress without the material preconditions that makes them possible is likely to end as well for us as it did for

Condorcet.

### III. The Other Gods

In Lovecraft's story *The Other Gods*, Barzai the Wise has learned so much and so long that he thinks he's the equal - no, the superior - of the Gods of the Earth, and doesn't need to heed their taboos. So he decides to sneak into one of their forbidden celebrations on top of a mountain and discovers, too late and to his horror, that there are Other Gods, the Gods of the Cosmic Process, "the Gods of the Outer Hells that guard the weak and feeble gods of the Earth". He comes to a grisly end.

This is the Suicide of Reason. It happens when the Ethical Process has succeeded for so long and so well that people have forgotten the Cosmic Process even exists, when we come to assume that everyone is naturally a rational actor, and forget that for most of human history, people have been tribal actors.

"Whereas the rational actor asks himself 'What is best for me?', the tribal actor must ask himself, 'What is best for us?' What matters for the tribal actor is not the pursuit of his enlightened self-interest but rather the success of his tribe. [...] The tribal actor [...] cannot take a moral stance outside the perspective of his tribe. For the tribal actor, the highest ethical idea is: 'My tribe, right or wrong.' The mere idea that his tribe could be wrong is unthinkable for the tribal actor, since he defines as right whatever the tribe deems as right, and wrong whatever the tribe deems as wrong."

Here comes the shaming code again:

"[W]hat limits his freedom is not so much the pressure of the tribal mind applied externally, but rather the fact that the tribal actor thinks with the tribal mind, and so cannot even imagine doing things differently from the way they are done by his tribe. For the tribal actor, departing from the ways of the tribe is simply unthinkable: He must remain true to his tribe through thick and thin."

The ultimate incarnation of the tribal actor is the fanatic, someone willing to do anything, even die, in the service of his tribe, and 'his tribe' can mean an actual tribe, a nation, a political movement, a religion, a race or whatever.

It's easy to see this as "Rational Actor: good. Tribal Fanatic: bad", but Lee Harries tries to get us away from that, because it obscures the fact that both modes can be good or bad. Since the nature of the fanatic is that he is willing to risk his life and die for his cause, it applies equally to Joseph Goebbels, determined to kill his family and himself rather than abandon Hitler, and to John Brown, leading his insurrection against slavery. A fanatic can even be a pacifist, as were the Buddhist monks who burned themselves alive to protest the Vietnam war, as was Mohandas

Gandhi, when he counselled Britons to willingly walk into the concentration camps rather than taint their souls with the evil of violence.

“The fanatic may be a saint or a terrorist, a revolutionary or a lone madman, while the rational actor may be a kind-hearted accountant, a devious business tycoon, a great scientist, a penny-wise house-wife, or an officious government bureaucrat.”

If you really want to grasp this, it's that the tribal actor alone can be a hero. When we call people heroes, we talk about soldiers, firefighters, front-line medical workers, people who are risking themselves. We never say it about people who are just maximizing their rational self-interest, like lawyers or stockbrokers, regardless of how nice they may be. To go back to the story at the start of this review, what would you consider morally better – to act tribally and unite to defend Scott from the predator, or to pursue our rational self-interest and run like hell? Even Ayn Rand, the high priestess of rational self-interest, depicted her heroic businessmen as basically fanatics, willing to endure just about anything in service to their work, and the number one criticism thrown at Objectivists is that we are too fanatical.

The problem with rational self-interest is that it can counsel you to do a lot of things, but never to become a hero, the kind of hero Nelson Mandela was when he declared that a free, democratic society was “an ideal for which I am prepared to die.”

“The rational actor is someone whose conduct is guided solely by his own enlightened self-interest, which, because it is enlightened, is willing to accept the rule of law. However, he is unwilling to die for anything, since death can never be in his self-interest, enlighten it however you please.”

It's difficult to argue with this point; when I look inside and consider those things I'm willing to die for – my family, for example – that knowledge has a different feel from the knowledge that I will sensibly put aside money for a rainy day each month.

The problem is that the rational actor can only exist in a world where the law of the jungle has been cleared away, because in the jungle, the rational actor is lunch. When the law of the jungle returns, the rational actor's survival relies on the tribal actor of his own tribe, even if he doesn't much like the tribal actor, even if he spends most of his time trying to break the tribal actor down. Since the rational actor isn't the product of reason, but the result of a shaming code that makes him ashamed to act unreasonably, he will not only oppose, but try to deny the reality of tribal fanaticism. The rational actor will oppose the tribal actor's return at all cost, even when the tribal actor is all that's protecting the rational actor.

This is one of those ideas, like O-ring problems or the motte-and-bailey fallacy, that once you know, you see everywhere. In 1941, with the Nazis and their allies having conquered pretty much all of continental Europe, H.G. Wells was arguing in print that the whole Nazi threat was ridiculously overblown, the ‘screaming little defective’ in Berlin couldn't really cause any serious trouble, the Second World War was a complete over-reaction, and what people needed to do is

be sensible: organize all civilised people into a big world-state that will deal with Hitler the same way any government deals with any criminal, and then we can all progress into the glorious future of free love, good hygiene and sane politics organised by reasonable men. Orwell commented as follows:

“All sensible men for decades past have been substantially in agreement with what Mr Wells says; but the sensible men have no power and, in too many cases, no disposition to sacrifice themselves. Hitler is a criminal lunatic, and Hitler has an army of millions of men, aeroplanes in thousands, tanks in tens of thousands. For his sake a great nation has been willing to overwork itself for six years, and then fight for two years more, whereas for the common-sense, essentially hedonistic world-view which Mr Wells puts forward, hardly a human creature is willing to shed a pint of blood. Before you can talk of world reconstruction, or even of peace, you have got to eliminate Hitler, which means bringing into being a dynamic not necessarily the same as that of the Nazis, but probably quite as unacceptable to ‘enlightened and hedonistic people’. [...] The energy that actually shapes the world springs from emotions – racial pride, leader-worship, religious belief, love of war – which liberal intellectuals mechanically write off as anachronisms, and which they have usually destroyed so completely in themselves as to have lost all power of action.”

I had my own encounter with exactly this mindset when I wrote a post for LessWrong. I criticised Eliezer Yudkowsky’s argument that Al Qaeda couldn’t be motivated by a hatred for freedom because people do not cast themselves as villains in their own stories<sup>7</sup>. What was weird wasn’t that I ran into people denying the reality of Islamic fanaticism but ones denying the reality of Nazi fanaticism – that Hitler couldn’t really have had the extermination of all Jews as his goal because that would be economically inefficient etc. etc. This weirded me out at the time, and to be honest, it still does: you had eminently reasonable people denying the reality of the Nazi ambitions as avidly as any member of the AltRight. Yet it makes sense under this scheme: here you have people utterly, totally, almost fanatically, committed to Reason, and this blinds them to the fact that other people might not share the same commitment.

Harris argues that today this sort of LessWrong-style epistemic failure isn’t limited to obscure grey tribe internet sects, it’s the norm throughout our society, no matter what politics we subscribe to.

“A carpe diem society has emerged in the liberal West today, and like all sophisticated societies, it has produced a self-serving ideology to convince itself of its own rightness. Curiously, this ideology is shared across our so-called political spectrum. Its basic tenet is that feeling good about yourself is the highest aim in life. Born-again Christians don’t worry about original sin; they’re saved, and they love Jesus because he makes them feel good about themselves. Liberal educators teach children that the highest virtue is self-esteem: thinking that you’re okay just the way you are. Many leftists today fashion their politics on the basis of what makes them feel good about themselves – they adopt causes that make them feel virtuous, enlightened, and superior, engaging in what Marx correctly derided as ‘utopian socialism’. Libertarians argue that the highest good is to follow your bliss, as Joseph Campbell put it. In many ways the maxim of

all carpe diem societies is best expressed by the popular song: 'Don't worry. Be happy.'

All of which brings us to the main point of the book, where it stops being generally and historically depressing and alarming and becomes specifically and contemporaneously depressing and alarming, the subtitle: "Radical Islam's threat to the West".

## IV. The Crash of Civilization

Harris starts with 9/11. American reactions to 9/11 were unlike any reaction in history – the first thing did George W. Bush did was to visit a mosque and proclaim that Islam was a religion of peace. The wars that followed, in Afghanistan and Iraq, were not punitive expeditions aimed at punishing an enemy, the way that the British air force flattened German cities in retaliation to the bombardment of their own; they were an attempt to solve the problem of Islamic fanaticism by bringing democracy to the Middle East – i.e. by bringing in the politics and culture of the rational actor.

Harris notes that there was no disagreement across the political spectrum that Islamic fanaticism could only be explained by root causes. Noam Chomsky and the left said the root cause was American foreign policy; Paul Wolfowitz and the neoconservatives said the root cause were the corrupt and dictatorial regimes of the Middle East. The neoconservatives won that 'argument' in the sense that they directed foreign policy, but the fundamental assumption, that fanaticism could only be caused by bad root causes and not be a power in itself, was never challenged.

This isn't often admitted because Bush's embrace of neoconservatism placed the Left in a quandary. Whatever you may think about the idea of overthrowing tyrants and spreading democratic revolution, conservative isn't it. At least in terms of foreign policy, Bush and the neocons moved so far to the left that they could be criticised by Francis Fukuyama as "Leninist". The Left needs to be to the Left of the Right, but if the Right is this far Left, what, as it were, is Left?

The result was "Bush derangement syndrome", the desperate attempt to insist that Bush and the neoconservatives were motivated by something other than their stated goals. It didn't really matter what, stupidity or greed or fundamentalist Christianity, but anything rather than admit that he was sincere about his idea of spreading democracy and revolution, that his real faith wasn't the faith of Christ, but the faith of Condorcet – that he was sincere in his belief that 'freedom is the desire in every human heart'.

I remember those days and those arguments and I can tell you that this faith was utterly mainstream, and it wasn't just shared by Republicans or Bush acolytes. Fareed Zakaria, a staple of CNN, wrote a book (The Future of Freedom) in which he argued that Iraq and Iran were ready for democratic revolution, and he supported the removal of Saddam Hussain. So too did the 'Reds for Bush' – those long time liberals and leftists like David Aaronovich, Nick

Cohen and most famously Christopher Hitchens, who decided that supporting Bush was a price worth paying for ridding the world of someone like Saddam. Hitchens always argued that this was a matter of solidarity with his socialist comrades in Iraq, and exporting revolution was what good Marxist should do. When he was attacked as a sell-out by his former comrades, he jeered that they had become complacent and conservative.

People who want to blame the failure in Iraq on Bush's venality or stupidity can't explain why the Obama administration faced the same fiasco in Libya – or why the 'Arab Spring' empowered Islamic reactionaries. Placing all blame for the failure in Iraq on the shoulders of Bush is a way of evading the magnitude of the disaster. If the failure is simply the defeat of one man's folly, it's no big deal; if it's a defeat of our whole worldview, that's another matter. If Iraq was the defeat of George Bush, you can relax; if Iraq was the defeat of Condorcet, then we are in deep trouble.

The 'root cause' argument, whether it's the root-cause of foreign policy or the root cause of lousy regimes, leaves too much unexplained. It can't explain why the post-Taliban regime in Afghanistan was united on the importance of killing Abdul Raman for the crime of converting to Christianity, nor why Muslims in Nigeria lynched Christians in response to cartoons published in Denmark. It can't explain why the majority of the victims of jihad terrorism are other Muslims, killed for being the wrong sect or not being devout enough. It can't explain the genocide of Christians in the Sudan or Hindus in Pakistan or Yazidi in Iraq. It can't explain why there are ten countries that punish homosexuality with death, and thirteen that do the same thing with atheism. It can't explain the widespread opposition to free speech, or freedom of religion or sexuality, not just in the Middle East but among Muslim communities in the West. Above all, it can't explain why the hope that Islamic fanaticism would give way before modernity has been proven wrong for at least two centuries.

"Not quite two centuries ago, when the English scholar of Arabic E.W. Lane first came to Egypt, he predicted that contact with European civilization 'will, probably, in the course of time, materially diminish the [Muslim] feeling of fanatical intolerance.' Yet in the final edition of his book, *Modern Egyptians*, Lane was forced to add in a melancholy footnote that his original 'prediction has not yet been fulfilled; on the contrary, European innovations in the dress and domestic manners and customs of the *grande*es, and of persons in the employ of the government, have enormously increased the fanaticism of those who belong to the religious and learned profession, and generally speaking, in the bulk of the population.' In short, contact with Western culture had not only failed to modernize the bulk of the Muslim population, it had actually made them more fanatically intolerant of Western ways than they were before."

While none of this can be explained by root causes, it can all be explained by the idea that the Islamic world has a popular culture of fanaticism every bit as strong as our popular culture of reason. This is where Harris differs from Huxley. Huxley took it as axiomatic that the Ethical Process would result in us becoming more reasonable – as we are drawn into larger and larger communities, we have to deal with each other reasonably. Smaller communities that refuse to do so will be squashed flat if they don't get with the program. Yet a culture of popular fanaticism can also create large-scale social cooperation, and that is what Islamic fanaticism has done –



create sections of the world (Egypt, Syria) so thoroughly Islamized that it's hard to imagine them as anything else.

Most writers in this genre – Robert Spencer, for example – tend to write this as “Islam bad, Christianity good”, or, anyway, “Muslims bad, Infidels good”. Harris is too smart for that:

“[T]here is nothing uniquely Muslim in [fanaticism]. In the middle of the nineteenth century, the bloody Sepoy rebellion was set off in India among Hindus over a completely unfounded rumor that the English had deliberately used the grease from cows in the casing of the bullets they provided to the native army, the Sepoys. Muslims also rebelled, because the rumor added that the grease from pigs had been combined with the grease from cows – but here the Muslims were only doing what other religious fanatics had done since time immemorial [...] [T]he Catholic Church had no problem condemning to death heretics who departed from the orthodox faith, and John Calvin in Geneva had few qualms about burning Miguel Servetus because of the latter's questionable theological notions about the trinity. [...] Fanatical intolerance, when undertaken to protect the existing cultural and religious traditions of a people, is thus by no means a phenomenon unique to Muslims. It is simply that the Muslims still retain this cultural defence mechanism in a world in which fanatical intolerance is no longer the practice of the other great religions.”

Nor is Iraq the first liberal nation-building project to come undone because of extreme fanaticism.

“[W]hen Napoleon ‘liberated’ Spain in 1808 from the corrupt and imbecilic Bourbon monarchy, he had no qualms about imposing the superior values of the French Revolution on the Spanish people. The rights of man and freedom of conscience were to replace Bourbon absolutism and the infamous Spanish Inquisition. Napoleon put his own brother, the humane and decent Joseph Bonaparte, on the vacated throne of Spain. With the best of intentions, King Jose, as Joseph was dubbed, set about freeing the Spanish people from the old iniquitous system that had kept Spain in the Dark Ages. He was aided in his efforts by many educated and enlightened Spaniards, who saw in the new king the answer to their prayers. Yet the Spanish people did not welcome the liberation from despotism and obscurantism that they had been offered by the French arm. In one of the few genuinely populist rebellions in European history, they spontaneously rose up against the new order and demanded a return to the old. They wished to put the vile and despicable Ferdinand VIII back on the Bourbon throne as an absolute monarch, and even clamoured for a return of the Inquisition. They developed guerrilla warfare, attacked the French troops, and even secured one astonishing victory against them at the battle of Baylen on July 20, 1808.”

The popular culture of fanaticism in the Islamic world means that the attempts to export democracy, whether it's Bush's project in Iraq or Obama's project in Libya or the Arab Spring, could have only one result. Free and fair elections didn't open the door to liberty, but to populism.

“American foreign policy makers, despite their noble intentions, made the fatal error of confusing populism with liberal democracy, thereby overlooking the fact that populism has almost invariably been the enemy of liberal cultures of reason, not only in the Muslim world, but in the West. Populism is the politics of the tribal mind, whereas liberalism is the politics of the rational actor.”

If this is ringing any bells, Harris beat you to it.

“The danger here is that the established political systems of the United States and Western Europe will face the kind of crises of confidence that lead to the dissolution of the rule by parliamentary democracy in Europe between the end of the First World War and the beginning of the Second World War, accompanied by a search for new and charismatic leaders who will emerge not from the established political parties, but from among the people themselves. These populist crusaders, unlike the established leadership, will be able to speak directly to the hearts of the people because they feel the same range and anger and betrayal that the people themselves feel. They will share their tribal code. For this is the essence of all populist leadership – the leader must be seen as someone who is loyal first and foremost to the people, and not to some higher abstract idea, however noble. They must be willing to act like the chief of a tribe, and not like the chair of a board of a multinational corporation.”

Ladies and gentlemen, President Trump. It's always worth remembering that, before he won the Presidency, Trump took on and wiped out both the republican establishment and the conservative movement. The landmark publication of official conservatism, National Review, dedicated a whole issue to “conservatives against Trump.” Fox News started out harshly critical of him, and the Republican Party made it clear from the start it didn't want him. Trump didn't care, and was happy to boast that he wasn't a “real conservative”; he trashed fundamental conservative principles, such as the free market, and yet he won, simply because he was the only one on the stage who was playing by the rules of populism.

Four years later, despite the global pandemic, the consequent recession, and his utter incompetence being clear for all to see, he still came within a hair of winning re-election.

“The struggle for survival will again create the Us versus Them mind-set to success in a life-and-death struggle, and in such an environment, the liberal internationalist will increasingly be looked upon as a traitor to his tribe. Nor does it make any difference whether the liberal internationalist is pushing the globalist agenda of multinational capitalist corporations, or the cosmopolitan agenda of the multicultural Left. From the point of view of the tribal mind, both forms of internationalism will represent a betrayal of the interests of the tribe.”

I'm old enough to remember the days when the Right was advocated global free trade. Watching 'globalism' become a curse word has been a shock.

This also explains the sheer weirdness of Trump's support. It's possible to make an argument that an unknown like Trump was a better bet than the known corruption of the Clintons, lesser of

two evils and all that. Yet the true MAGAists showed a loyalty to Trump that makes no sense by any rational standard. It does, however, make sense if you think of Trump not as a politician chosen to do a certain job, but as a tribal chief. You defend your chief no matter what, because if the chief falls, the tribe goes with him.

It also explains why there's been a steady stream of voices who hated Bush, called him every name in the book, suddenly starting to pine for him – all those articles about 'crying wolf'. Bush was a known quantity who played the game by the standard rules. Trump just kicked the board over and said "Why should I play by your rules? You can all play by mine."

What's more, Trump is hardly the only populist to be empowered. In Eastern Europe, almost all governments are headed by populists, while in the West, they're closer to taking power than at any time since 1945. Most significantly, the world's largest democracy, India, is led by the populist leader Narendra Modi.

Continuing with his habit of saying "That writer who is notoriously pessimistic and depressing? Not nearly pessimistic and depressing enough.", Harris argues that Samuel Huntington's Clash of Civilizations gets it wrong. A clash is when two power blocks rub up against each other and try to get the better of each other. China and America each try to expand their zone of influence, for example. One or the other has dominance over, say, Taiwan, but no one expects American troops to land in Shanghai or Chinese troops to seize San Francisco. The danger we're facing isn't a clash of civilizations, but a crash – that Islamic radicals will do to the Western order what Trump did to the American political status quo.

It's in their best interest to do so. Under a clash scenario, Islamic radicals are doomed so obviously that even they know it. Imagine the Muslim Brotherhood coming to power in Egypt and using the Egyptian Air Force to attack America. They wouldn't last a week.

Yet what's actually happening – terror strikes launched by stateless enemies who hide within the larger community – is a threat that our systems and leaders have no way to deal with, or, rather, no way that's morally acceptable to us. In his withering critique of Victor Davis Hanson, the War Nerd noted that Hanson's appeal to classical models of warfare leaves out the fact that every ancient power that faced the kind of civilian insurgency that America faced in Iraq would have responded by wiping out the civilian population<sup>8</sup>, something utterly unacceptable to everyone across the political spectrum today. So the only answer available to our leaders is the one that our leaders are already doing: ramping up policing, trying to prevent terror attacks before they happen, and try to hold together the politics of reason as best they can, by finding compromises and accommodation where possible, while reinforcing the shaming code that shames people from acting unreasonably, which means ramping up the political correctness.

Unfortunately, this combination of trying to find an accommodation with Islamists through compromise while at the same time trying to shame the native population through political correctness is absolutely guaranteed to feed the populist fury that men like Trump exploit. People are not interested in high minded lectures on tolerance from those making

accommodation with the most intolerant movements on earth. The most ardent Obama supporter must have facepalmed when, following the Orlando nightclub massacre, he declined to even use the words 'radical Islam' and Clinton tried to make the issue about gun control. Trump could not have asked for a better gift.

There isn't any sign this is going to get better, and a good number of signs that it will get worse. Trump is a dishonest, incompetent crook, and that's a good thing. Imagine what he'd be capable of if he were competent, if he actually believed in something. Trump won because he was the only one playing the populist game. That will not be the case next time around.

Everyone who wants to be President will have learned from his victory, which means that the next competition will be between several different populists, with victory going to the one who really, truly means it: who is willing to build a wall, jail his opponents, ban all Muslims from entering the country, support mass deportations. And even that will be a kindergarten compared to what might happen elsewhere; the last time a populist leader rose in Europe driven by anti-Muslim animus he started throwing them into camps, and it took the combined might of American and Britain to unseat him.

Towards the end of his book, Harris writes four of the most haunting sentences I have ever read:

"Will the devastatingly effective institution of jihad, adapted to new circumstances, be able to transform the entire world into the realm of Dar al-Islam? No. But it may well be able to destroy the world that Western liberalism has made. And that is something to think about."

## V. How screwed are we?

I often think that Lee Harris is the smartest guy you've never heard of. He's read all those books you mean to get round to one day, and all the authors people namedrop (how many people talking about 'cultural Marxism' have actually read Marx?). That he can make the notoriously difficult Hegel understandable and even plausible tells you something. Yet he's weirdly obscure.

A gay guy who talks about a threat from radical Islam and American exceptionalism should be a regular on Tucker Carlson. He isn't, and when you read the book you can see why: America is exceptional, but that exceptionalism could only come about because of the genocide of native Americans? America's founders weren't smarter or better, just luckier? Bush was being traduced by the left, and he was sincere in his goals, which were the liberal internationalism of Condorcet, Kant and Karl Marx? American red tribe evangelicals are just as self-flattering as blue-state SJWS? We're all the spiritual descendants of a French intellectual? Political correctness is an attempt to preserve the best traditions of the West? The populism that put Trump in the White House is the Voice of the People and that is why it is a Very Bad Thing?

Huh?

Harris is impossible to pin down politically. If he's a liberal, he's a liberal in the sense we all are, and if he's conservative, he's the kind of conservative that'd be just as opposed to eliminating the welfare state as to single payer healthcare. Right, left, liberal, conservative: the labels don't apply.

The Suicide of Reason is a frightening book, not in the book review sense, but in the staring-at-the-ceiling-wondering-about-the-future sense. The only thing I have read that gives me comparable willies is Scott's discussion of AI risk. So: is it true?

Well, when someone comes up with a carefully argued theory that explains a great deal that's been previously frustrating and incomprehensible, and then uses that theory to make specific predictions that all come true, I tend to think he's on to something. The Shaming Code is an idea that explains so much you'd be foolish to dismiss it. That human beings are fundamentally tribal actors, rather than rational actors, is clear from a quick glance at human history. It's also a valuable epistemic tool – that anything you think is completely natural and normal is likely just your shaming code talking, and you should scrutinize it ten times as much as any other proposition.

That's something I know from a personal experience that I'm grateful for but not very proud of. Modern liberals tend to lump all bad things – racism, sexism, homophobia etc. together. A racist is likely a sexist is likely a homophobe and so on. But where I grew up, racism was taboo, as was homosexuality. So I was both anti-racist and homophobic until about fifteen, when I moved to a much different society, discovered one of my closest friends was gay, and did some rethinking. As I say, I'm not proud of it, but I am grateful for the experience of thinking something is utterly normal and natural and yet being completely wrong.

That rational actors are driven out by fanatics in cultural struggle for existence is also pretty clear – following the Danish cartoon riots, many asked why all newspaper publishers didn't just republish the cartoons at once, demonstrating that they weren't going to be intimidated? Well, because they were being rational actors, decreasing the threat to their own persons as much as possible. To publish and be damned, to risk their lives for a principle, would have required them to think not as rational businesspeople but as something closer to Christian martyrs or Marxist revolutionaries, i.e., as fanatics.

My main point of disagreement with Harris is the slight elision between fanatic and tribal actor. Going through the book again, it's clear to me that Harris defines a tribal actor as someone who is utterly committed to the tribe, to the extent that he cannot see any good beyond the tribe's good. A fanatic on the other hand, is someone willing to die, or at least risk death, for something beyond himself. These aren't the same thing. Whatever you think about Gandhi insisting that suicide is better than violence, it can't be compared to the fanaticism of Joseph Goebbels. There is a fundamental difference between Umkonto we Sizwe, the militant arm of the ANC, and HAMAS. Conversely, the carpe diem societies of the developed West have had

no problem recruiting and training devastatingly effective soldiers who routinely risk their lives but would balk at the indiscriminate massacre of enemy civilians.

To be fair to Harris, he's not denying any of this. It's merely that, in a time and place that regards our popular culture of reason as a normal thing that will just spread automatically expand to cover the world he is shouting, as loudly as he can, "GUYS, IT'S NOT SO, OUR CULTURE IS REALLY, REALLY RARE, AND WE CAN LOSE IT QUITE EASILY"

The problem is that it is the scariest parts of his book that are the soundest. Iraq, Libya and the Arab Spring have all failed in exactly the way Harris predicted. We are facing a wave of populism that has just the characteristics he defined. One of the most worrying things in the Trump campaign was when his then spokeswoman dismissed concerns about the Muslim ban with "So what? They're Muslims!" That's the tribal actor at its most basic – what happens to members of the other tribe isn't our concern.

Ideas about a Eurabian caliphate have always been fanciful. On the other hand, the idea that Islamic fanaticism manages to crash our civilization is all too believable. Historically, Islamic jihad has tended to call forth reactions at least as savage – the Crusaders, Hulaghu Khan, Vlad the Impaler or Slobodan Milosevic. There would be nothing to laugh about in a similar reaction empowered by twenty-first century weaponry.

Harris offers two ways out of the current mess, which he calls enlightened tribalism and critical liberalism. "Enlightened tribalism" is like those northern American states that regarded slavery as evil and completely banned it within their borders, but did not want to abolish slavery worldwide. "Critical liberalism" is like the radical abolitionists who wanted to abolish slavery worldwide. In our day, both would agree that our rare cultures of reason are ethically superior to those where fanaticism and violence are the norm. Both would accept that our cultures of reason are the result of a unique visceral code. Most important, both would accept that there are points where being a rational actor is not sufficient, where you have to fanatically defend those values that make the rational actor possible in the first place – being as prepared to die for them as Condorcet in Bourg-la-Reine or Mandela at the Rivonia Trial. The main difference would be that the enlightened tribalist would be content to defend these values within the bounds of the Western world and sharply reduce or eliminate immigration, while the critical liberal would support any movement worldwide that tried to spread these values.

The idea that you could fanatically defend the values of the rational actor appears paradoxical, but that's because of the aforementioned elision between the rational actor and capital-R Reason, in the Enlightenment sense. A rational actor is someone acting for his rational self-interest and therefore can never be brought to sacrifice himself, nor really put himself at risk. But capital-R Reason has had no problems recruiting fanatics, zealots and martyrs, and these have succeeded far beyond anything purely tribal actors ever have. Napoleon nearly conquered all of Europe, and even after his defeat managed a return to power, based on the conviction he inspired, which was based on the belief in Enlightenment Reason. Similarly, we can understand the historical accidents that allowed the Enlightenment to get underway, but

once it was underway, it seems to have spread exactly as Condorcet hoped. Harris writes that Condorcet was naïve to assume that all that was needed was to show how well societies like France and America and England were succeeding and explain to others how they could succeed this way – and yet Condorcet’s ideas have spread far beyond those three nations, not merely in Europe or ‘the West’ but globally. Sit down with the average Kenyan or Ghanaian and ask him what he would like his government and society to be like, and I guarantee you will hear something Condorcet would have approved of.

The tide of reason even seems to be even eroding Islamic fanaticism. Atheism is spreading in the Muslim world, to the extent that now 5% of Saudis say – privately, anonymously – that they are atheists<sup>9</sup>. Other movements, like those for women’s rights, are more pronounced. These are driven by the same thing that led to tolerance in Europe – the sheer horror of religious violence. As the Thirty Years’ War paved the way for the enlightenment in Europe, ISIS and its related movements seem to be doing the same thing in the Middle East.

It has not escaped me that both Enlightened Tribalism and Critical Liberalism will prove hard sells in today’s political landscape where the very idea of better cultures is taboo. But perhaps they will be more easily accepted as people understand that the alternative is a process of crash and decivilization that ends with the Muslim world a smoking, radioactive waste, and Muslims within the Infidel world staring out through barbed wire – as they are already doing in China.

---

1 <https://www.psychologytoday.com/us/blog/minority-report/201406/asian-honor-and-suicide>

2 [https://www.ted.com/talks/steven\\_pinker\\_the\\_surprising\\_decline\\_in\\_violence](https://www.ted.com/talks/steven_pinker_the_surprising_decline_in_violence)

3 <https://www.newsweek.com/racism-madison-high-school-teacher-fired-n-word-1466175>

4 <https://arxiv.org/abs/2001.09318>

5 [https://wiki.lesswrong.com/wiki/How\\_To\\_Actually\\_Change\\_Your\\_Mind](https://wiki.lesswrong.com/wiki/How_To_Actually_Change_Your_Mind)

6 <https://slatestarcodex.com/2019/10/16/is-enlightenment-compatible-with-sex-scandals/>

7 <https://www.lesswrong.com/posts/qpuufdn37Kav8jZcN/if-you-can-see-the-box-you-can-open-the-box>

8 <https://www.theamericanconservative.com/articles/its-all-greek-to-victor-davis-hanson/>

9 <https://www.washingtontimes.com/news/2017/aug/1/atheists-in-muslim-world-growing-silent-minority/>

# The True Believer: Thoughts on the Nature of Mass Movements by Eric Hoffer

## I.

Eric Hoffer sets out to create a unified theory of mass movements. And he doesn't just mean to tell you about some patterns in the development of political movements. He means to say that whether you are joining a political, a nationalist, a religious or a revolutionary movement, you are probably doing it for the same reasons, and you are probably going to play a role in a remake of the same old play. The premise of [The True Believer](#) seems destined to end in cherry-picked examples from history, further cherry-picked for characteristics that match some template. And if I were to imagine someone who might have a broad enough knowledge base and the qualification to do rigorous enough analysis to calm some of my concerns, Hoffer is the opposite. If his own claims about his education are to be believed, he did all of his learning while dividing his free time after migrant work between brothels and a library.

And yet he mostly pulls it off. The book repeatedly gives you jolts of insight; and Hillary Clinton is still telling people working on her campaign seventy years after its writing that they need to read it to understand contemporary America.

Hoffer's most original argument is that there is a fixed pool of people who are in the right kind of situation to be open to joining a mass movement. So a leader is competing for the same audience whether he is recruiting for a religion, a fascist party, or a communist revolution. This lines up surprisingly well with Jonathan Haidt's recent moral foundations [research](#), where he finds that people on both the far left and far right, as well as religious people rely heavily on purity as a moral foundation. Hoffer's description of a true believer is more determined by his immediate circumstances than Haidt's would be, but Hoffer seems to have glimpsed the underlying structure of moral motivation decades early. He describes how both fascist and communist leaders in the first half of the twentieth century would boast about how good they were at poaching followers from each other. These potential converts are people who feel that their lives are spoiled, and that they can't individually fix anything. They want to subsume themselves in something bigger, and forget about their individual lives. The largest group are the new poor, who have recently lost what they had come to expect. His second group are misfits, who haven't found their place in life. These range from unemployed college graduates, to veterans, to new immigrants to the irreparably defective. These are followed by a grab bag of other outliers, from the inordinately selfish, to minorities in the process of losing their traditional culture, to bored spinsters, to criminals.



The second argument of the book is that movements have a strange mood. Everyone is frustrated, but no one actually wants things to get better in their lives right now. The true believers are primed for self-sacrifice, for their lives to actually get more difficult in the present, all for some idealized distant future. And in their eagerness to abandon the self, they all rush into a world of make-believe, where they are cast as the heroes in a theatrical production, unafraid of death. Hoffer talks about how Hitler's costumes and rallies were directed towards this aim; but we could as easily talk about the live action role-play feel of Jan 6, 2021.

Hoffer's final big idea is that all movements have distinct stages, each led by a different type of personality. The first stage is led by "men of words with a grievance." They speak and write to criticize and undermine the existing order, priming the movement. This again seems to be an impressive premonition of future, more rigorous analysis. In *Ages of Discord*, Peter Turchin argues that cycles of unrest are set in motion by there being too many educated people with frustrated expectations of power at a given time. The second stage is the destruction of the status quo, led by fanatics. The fanatic leader is not interested in reform, but only in wholesale destruction and chaos. The final stage is the consolidation of a mature movement, led by "practical men of action." The fanatic is pushed aside as being too radical, and more predisposed to constant destruction than the building of a new order. At this point the movement stops arousing selfless devotion, and starts to maintain itself through institutions, propaganda and coercion.

## II.

Much of this sounds like it's getting at something deep, but does it actually describe the majority of movements? Or does this all seem original because Hoffer is describing counterintuitive features of some of the weirder movements of history. Let's try to spot check the generality of a couple of his more interesting arguments.

Are the "new poor," who are not *that* poor, but just downwardly mobile, really at the center of movements? Hoffer bases this generalization on all of the "ruined middle class" Germans and Italians who joined Nazi and Fascist movements. He even goes so far as to say that historically economic depressions didn't really frustrate the masses until enough people got wealthy enough to really have something to lose. Looking outside of his examples, the theory seems to do surprisingly well describing the 1905 Russian Revolution. The recent advancement of peasants from serfdom to land ownership, the economic depression around 1900, and the frustrated veterans returning from the Russo Japanese war, all seem to have played a role. As far as I can tell, the theory does worse with the French and Chinese communist revolutions, where the peasants seem to have been much poorer (and in the case of France, were maybe just starting on their way up). And it would probably do even worse with the first converts to Islam, who were apparently the poorest members of the surrounding society. These are the kinds of people labeled the abject poor by Hoffer, about whom he writes:

*To be engaged in a desperate struggle for food and shelter is to be wholly free from a sense of futility. The goals are concrete and immediate. Every meal is a fulfillment; to go to sleep on a full stomach is a triumph; and every windfall a miracle. What need could they have for "an inspiring super-individual goal which would give meaning and dignity to their lives?"*

Next let's turn to a few examples after Hoffer's time, and therefore a sort of out-of-sample test for his theory. The story appears to line up well with narratives of suicide bombers coming from the middle classes, and the more educated segments of society leading protests in Tiananmen Square, during the Arab Spring and during the Color Revolutions. However, even there, one has to be careful in that Tiananmen Square and some of the color revolutions happened during times of increasing prosperity and reform, rather than during impoverishment. The example that seems to push the balance of evidence in Hoffer's favor is the civil rights movement in 1960's America. It recruited heavily from those in the black population who had started to advance economically through the 40s and 50s, as well as the criminal population in the case of the nation of Islam. And at least some of the [precipitating events](#) appear to have been early automation, white flight, and relocation of urban factories leaving many of the millions who had moved to cities for work unemployed.

How about the claim that there are stages of all movements, characterized by men of words, followed by fanatics, followed by practical men? Hoffer offers Robespierre, Lenin, Mussolini, Hitler, Stalin, F.D.R. and Gandhi as some of his fanatics (although the latter two with qualifications). He is far less explicit in providing distinct examples of frustrated men of words and men of action, presumably because many of them don't make history. He also admits that many of his examples serve in more than one of the three roles.

If we take Malcolm X and a potential historical Jesus Christ as men of words, they seem to support the idea that at the beginning of movements, there are intelligent men excluded from the mainstream and critical of it. Malcolm X describes a teacher who told him he could never be a lawyer, and contributed to his choice to drop out of school and pursue a life of crime. Jesus criticizes the legal expert scribes and Pharisees (who apparently have a higher social standing than him), and the institutions they represent. Both of them also show that it is possible for there to be men of words even without them having to rely on the written word. Both operated at a time when many prophets and critics preached in public places to anyone who would listen. A more explicit supporting example that comes to mind is Anne Applebaum's [description](#) of her former friends who have risen to the top of Poland's new far right parties. She says many of them joined the liberal democratic anti-communist movement in the 1980s because they were frustrated that they weren't being recognized in that regime, and have now joined anti-liberal movements because they feel like western institutions haven't given them enough recognition or respect either. On the other hand, the founding fathers who wrote to prepare the way for the American revolution weren't at all frustrated in their social aspirations, they were at the top of society in the richest part of the United Kingdom. Similarly, the leadership of the Students for a Democratic Society (SDS), who would radicalize to become the Weather Underground, were all doing pretty well at elite universities. There are probably lesser known men of words criticizing

the order around the start of any movement, but it's not at all clear that they are actually a necessary ingredient rather than a side effect. The easier explanation seems to be that when people are frustrated enough to throw their life into a movement, there will be a few of them who are also predisposed to put it into words. If one of them is influential enough, he will often rise to the position of fanatic leader (as Hoffer admits about Lenin, Hitler and others). If he is not, it's not clear he plays much of a role aside from hardening his own belief in the process of preaching.

The label of Fanatic also seems to be informative, but not clearly applicable to most leaders of movements in their active phase. Hoffer's description of the will to totally destroy the current order seems to be built around the likes of Lenin and Hitler. But this is far less obviously true the founders of religions, even less so about leaders like F.D.R and Churchill, and even less about M.L.K.

I find the Practical Man of Action category less enlightening yet. It seems to be restating the obvious, that if a movement succeeds it will need to be maintained by someone. Hoffer also admits that this person will often just be the Fanatic serving a new role, and has no slam dunk examples of a fanatic leader being pushed aside by the practical men. However, Hoffer does have some interesting observations about the methods such leaders use to try to maintain the devotion of their followers. He implies that coercion, and keeping the masses reliant on the new leadership are essential, and goes on to predict:

*The most dangerous moment for the regime of the Politburo will be when a considerable improvement in the economic conditions of the Russian masses has been achieved and the iron totalitarian rule somewhat relaxed.*

This is not exactly a detailed prediction of how the Soviet Union would collapse forty years after he wrote the book, but it's enough to be a little eerie.

Overall, Hoffer comes out pretty well. His description of the mass of True Believers seems to point out something important and surprising and yet mostly true. I read a bit about the 1905 Russian revolution and the motivation of the civil rights March on Washington as I was doing my spot check, expecting to find that the participants had actually been on their way up economically. Instead I found Hoffer's "new poor" description to be surprisingly accurate, and to find other observations from the book unexpectedly confirmed. Hoffer's description of the types of men who lead each stage of a movement seem less insightful, and less clearly accurate, but it still provides a potentially useful lens.

### III.

What does this all say about the United States, and especially our current moment in politics? Hoffer talks about how America has few movements because it has had the outlets of westward

migration and self-advancement to prevent the buildup of frustration. We can add the individualism and optimism often highlighted in the American attitude as additional defenses. The relevance of optimism may be best captured by Martin Seligman's [claim](#) that the more optimistic candidate won US presidential elections the vast majority of the time in the 20th century. Americans do not seem to go for the oratorical style of deprecating the present. Hoffer claims is central to movements. And yet, Seligman admits that optimism stopped being a reliable predictor of elections around 1988, and that Trump was clearly the more pessimistic candidate in the 2016 election. Internal migration within the United States has plummeted over the past decades. Social mobility (and therefore self-advancement) has plummeted as well, to a level well below European levels. Perhaps we are not as immune to movements as we once were?

The United States was built by the frustrated immigrants who Hoffer admits may otherwise have joined a movement. Could this have left an imprint on our innate or cultural legacy leaving us more predisposed to movements, even if circumstances did not allow these tendencies to play out for the first couple of hundred years? And perhaps this predisposition is reflected in our penchant for violent crime unique in the developed world. In fact, Hoffer claims crime is often an alternate outlet for a potential true believer. Similarly, the level of religiosity in this country and our waves of religious revivals may reflect the same predisposition. And now we come to a moment in our history when for the first time the large majority of people can expect to be economically [worse off](#) than their parents. What's more, Hochschild [describes](#) how many now feel resentful that minorities, immigrants and women are "cutting in line," unfairly leaving them behind. Is it possible that our previous optimism and individual advancement only mean a bigger moment of reckoning now that those trends reverse? Maybe our deaths of despair (from opioids, alcohol and suicide), our polarization, our low trust in government are all just warning signs that our underlying predisposition and our recent reversal of fortune are about to create a moment overripe for movements?

Nothing today yet appears truly worthy of the name mass movement, where people sacrifice their comfort and lives for some distant goal. And yet much in the present rhymes with Hoffer's observations. Are the underemployed college graduates who supported Bernie Sanders, and conservative commentators excluded from liberal media and academia, versions of frustrated men of words? Is Qanon a manifestation of the self-contained and nonsensical nature of an idealized movement Hoffer describes in:

*Thus the effectiveness of a doctrine should not be judged by its profundity, sublimity or the validity of the truths it embodies, but by how thoroughly it insulates the individual from his self and the world as it is. What Pascal said of an effective religion is true of any effective doctrine: it must be "contrary to nature, to common sense and to pleasure."*

Or does this sound at all relevant to the current moment?

*Indeed, it is easier for the frustrated to detect their own imaginings and hear the echo of their own musings in impassioned double-talk and sonorous refrains than in precise words joined together with faultless logic.*

All of this has me concerned, especially when combined with Peter Turchin's prediction from ten years ago that the coming decade is going to be a tumultuous one. Turchin's theory has some of the same causal mechanisms as Hoffer's, and it's already had some confirmation with the storming of the capital. But at the same time my main intuition is that even if these theories are valid, we are living in a very different time from *The True Believer's* 1951. We are living in a world with social safety nets that make it harder to feel desperate, streaming TV to quiet the frustrated mind and stave off the boredom Hoffer describes as a vital ingredient of movements, and a growing distance from experiences of self-sacrifice in wars.

At the same time, maybe a movement is what America really needs right now? Hoffer says about the fanaticism of movements:

*And it is strange to think that in receiving this malady of the soul the world also received a miraculous instrument for raising societies and nations from the dead—an instrument of resurrection.*

Maybe America's anemic economic growth and decreasing effectiveness of government of the past few decades are signs that we are desperately in need of a resurrection? Maybe the antidote to the decadence some complain about today is a reinvigoration through a mass movement. The chaos of movements seems generally very dangerous for the short term prospects of a country, and often the long term prospects as well. And I am comforted to think that our modern world is one far less susceptible to truly totalizing movements. And yet there is the chance that without them our status quo is a slow one-way ratchet in the wrong direction.

## IV.

Now let's accept for a moment that Hoffer's mass movement is the generalized human response among the most frustrated at frustrating historical moments. And for a brief speculative overreach, let's ask, does that imply anything about more everyday existence? In the same way that mental disorders are often on a continuum, and seem to provide insight into all human minds, is a movement just an extreme version of the standard human response to frustration?

Hoffer mentions a proneness to hate and credulity as attributes of the true believer. We see watered down hate among people going through a difficult time, who fall into depression which manifests itself as constant anger. We see it in the increases in spousal and child abuse during recessions. Perhaps we even glimpse it in the spikes in race riots in the United States during the unrest of the early 20th century, and the increases in witch burnings and ethnic violence

during difficult economic times historically. Increased credulity might be what drives the suffering sick to quack medicines and the angry to conspiracy theories.

To consider a more encouraging attribute of true believers, we can look at Hoffer's "willingness to dissolve [the self] by losing one's individual distinctness in a compact collective whole."

Perhaps turning to alcohol, religion or meditation in times of trouble are all just variations on this dissolution of the self. Much has been said about suffering as a route to transcendence. Maybe Hoffer's "desire to escape the blemished self" happens at all moments of frustration, and at the extreme is what leads to most spiritual experiences.

# The True Creator of Everything: How the Human Brain Shaped the Universe as We Know It by Miguel Nicolelis

Miguel Nicolelis seems to be a pretty smart guy. He's the co-director of the Neuroengineering Center at Duke University, and has clearly forgotten more about superior longitudinal fasciculi and frontoparietal neuronal circuits than most people will learn in their entire lifetimes. He wrote a (seemingly decently received) book in 2011, *Beyond Boundaries: The New Neuroscience of Connecting Brains with Machines* (Times Books). Most famously, Nicolelis' work on brain-machine interfaces led to a remarkable scene at the opening ceremony of the 2014 World Cup, when a 29-year-old man paralyzed from the waist down kicked a soccer ball using a robotic exoskeleton.

Now, Nicolelis has written a new book, *The True Creator of Everything: How the Human Brain Shaped the Universe as We Know It* (Yale University Press, 2020). As titles go, this is, um, a bit bold, but I'm willing to chalk that up to the exigencies of the publishing industry—after all, it did prompt me to pluck this tome from among the four impressively solid racks of science books at my local Barnes & Noble. (I live in Raleigh, NC, and it's always encouraging to see a bookstore where the science section is more than a fraction of the size of the Bible/Christianity section.)

For this hopefully somewhat ambitious review of this very ambitious book, I'm going to try something different to help drive discussion in the comments thread. I don't claim to be a PhD-level expert in neuroscience, computational complexity, evolution, the origin of mathematics, or any of the other topics discussed in *The True Creator of Everything*. I am, however, fairly conversant and well-informed in these and other scientific topics, and have worked in science-adjacent fields for much of my career. I'm going to call out parts of this review where I think Nicolelis' conclusions are open to debate, based on my understanding (or lack of understanding) of the issue, and throw out specific questions to you, the readers.

## A Totally Organic Experience! (With a Tip of the Hat to Clairol)

How does the human brain process information and generate qualia? Nicolelis uses the concept of an "organic computer," which requires some unpacking. According to him, the brain turns "Shannon information," the streams of digital bits created by speech, touch, smells, and other sensory inputs, into "Godelian information," e.g., thoughts, memories, images, etc. that are embedded into brain tissue and thus strictly analog in nature. I will admit this all sounds a bit jargony—I'm still not quite sure why he calls analog information "Godelian"—but the idea gains force with repeated explication:

*Rather than being binary and digital, Godelian information is continuous or analog, given that its embedding in organic tissue is fueled by the process of energy dissipation to organisms. As such, Godelian information cannot be digitalized or discretized and treated as binary bits of information flowing in a noisy communication channel. The more complex an organism gets, the more Godelian information is laid out, embedded in the organic matter that forms it.*

Basically, I think what Nicolelis is saying amounts to a variant of “a picture is worth a thousand words.” Converting an (analog) mental image or feeling into (digital) words or other exterior representations would require an intractable amount of data. I think this gets pretty close to what most people informally mean by “qualia”: experiences that can only be conveyed approximately by words, painting, music, or any external creation of the human mind. To really understand a quale, you have to directly experience it, and that is only possible via the medium of your first-person-perspective human brain.

Nicolelis also lays great stress on the distributed, non-local, and extremely plastic functioning of the brain, all of which point to an analog rather than digital architecture. One surprising thing I learned from this book is that, according to experiments, different groups of neurons are recruited by the brain to perform the exact same motor functions: every time you raise your right index finger, for example, this action corresponds to the firing of similar (but not identical) groups of neurons. This seems very different from the way a digital computer would be expected to behave. Also, Nicolelis points out, specialized areas of the brain can be “retrained” (by experiment or necessity) to perform other functions, such as the visual neurons of blind people becoming more sensitive to the sense of touch. And, finally, decades of brain scans have shown that widely separated areas of the brain are often involved in the simplest acts and behaviors.

Why does it matter whether the brain processes and stores information in a digital or analog fashion? Well, you can probably see where this is going: Nicolelis is a strong disbeliever in strong AI. For one thing, as he points out, there is no hardware/software distinction in the human brain; he likens the brain to a tree, which converts exterior environmental information into (among other things) the rings that are an inherent part of its structure.

**[Question #1 for the SSC commentariat:** Does the hardware/software distinction, referenced above, really apply to today’s computer architecture? I have a hard time imagining researchers building a superdupercomputer and then loading up a beta copy of Microsoft Turing Test™. If this distinction between hardware and software isn’t as unique as Nicolelis thinks it is, then it would tell against his argument for the brain’s noncomputability, but perhaps not totally destroy it.]

In the most memorable line of this book, Nicolelis nails down the essential difference between a brain and a computer:

*You can’t reverse-engineer something that was never engineered in the first place.*

Here, Nicolelis is addressing the hope of strong AI researchers that mimicking the architecture of the human brain, using digital components, can bring a self-conscious AI into existence. He makes a strong case that this effort is doomed to failure. Whether you take the starting point as three billion years ago (the origin of life on earth), 400 million years ago (the origin of mammals), 70 million years ago (the origin of primates), three million years ago (the origin of human-like primates), or even 10,000 years ago (the origin of modern civilization), the human brain has been shaped by a near-infinite number of environmental and evolutionary influences over innumerable generations. What’s more, individual human brains are shaped by a lifetime’s worth of interactions with their environments, both experiential (parenting, social and political structures, etc.) and chemical (hormones, neurotransmitters, etc.).

There is simply no way, Nicolelis argues, to replicate these intractable quantities and arrangements of information in a reverse-engineered computer; something essential will always



be missing. He drives the point home with another telling analogy: a pool table can be considered almost infinitely simpler than a human brain. Yet, calculating the knock-on effects of a cue ball hitting a second ball, which hits a third ball, etc., by the time you get to the eighth ball you would have to factor in the gravitational field generated by the human player to effectively predicts its trajectory. A real-life analog system, whether it's a human brain or a pool table, cannot be effectively simulated by a digital computer.

**[Question #2:** Strangely, in his discussion of whether the brain is a digital or analog computer, Nicolelis does not reference the Penrose/Hameroff theory of quantum computing in brain tissue, though he does reference Penrose elsewhere. However, my somewhat naïve understanding is that a sufficiently powerful quantum computer can model an analog system. Can Nicolelis' argument that the brain is an analog computer be reconciled with the theory that it's also a quantum computer?]

Nicolelis does offer a loophole of sorts for supporters of strong AI. One claim he makes, which I am not in a position to judge, is that highly organized patches of white matter in the brain act as "biological solenoids," emitting electromagnetic fields that synchronize brain activity and the firing of neurons. (He backs up this claim, at least partially, with experimental evidence that transcranial magnetic stimulation, or TMS, can help allay some of the effects of mental illness. I will leave it to our host, Scott, to weigh in on whether this is actually a useful and generally accepted technique.) If a new kind of digital computer architecture can incorporate a magnetic "globalizing" component, Nicolelis implies, it may come closer to duplicating the functionality of a human brain.

Ironically, towards the end of his book, Nicolelis speculates that our current immersion in digital culture (smartphones, virtual reality, the internet, etc.) might actually wind up degrading the analog functioning of our brains. He cites as an example old-school London taxi drivers, MRI studies of whom reveal unusually large hippocampi; what, he wonders, will happen to the hippocampi of drivers who learn to navigate London's streets not by experience, but by using a GPS system? I myself have wondered about the long-term effects Google will have on human grey matter; after all, there's no need to memorize (for example) how many Academy Awards a given movie won, in what year, and in which categories, when you can just look that information up instantly, online.

Nicolelis, who is clearly not afraid to coin a phrase, calls this the "digital chameleon hypothesis": as modern social structures and workplaces encourage and reward responses to digital inputs, the human brain responds by "pretending" to be digital, and loses many of the analog qualities that made us human in the first place.

Anyway, all of this is just the first part of Nicolelis' book. On to our next section, brainets!

## Check out the Big Brainet on Brad!

There is a story, possibly apocryphal, about an abysmally amateurish production of the play *The Diary of Anne Frank*. Early in the action, an SS man enters stage left, whereupon a member of the audience shouts, "She's in the attic!"

That's how I felt reading Nicolelis's lengthy chapter about what he terms, rather unfortunately, "brainets" (which somehow calls to mind both Skynet and Raisinets):

*Basically, a brainet is a distributed organic computer composed of multiple individual brains that become synchronized—in the analog domain—by an external signal such as light, sound, language, chemicals, or radio or electromagnetic waves and, as a result, is capable of producing emergent collective social behaviors.*

Does any of this sound familiar? Well, let's simplify a bit—instead of “radio or electromagnetic waves,” let's just say the radio broadcast of *War of the Worlds* or the 5,000-screen release of *Avengers: Endgame*. Instead of light, sound and language, let's say that *Triumph of the Will* rally choreographed by Leni Riefenstahl. Basically, Nicolelis seems to have rediscovered the concept of the meme—which is kind of strange, since he references Richard Dawkins and memes, in passing, later on in the book.

Interpreting this concept as charitably as possible, Nicolelis aspires to add a new layer of meaning to the meme regarding the existence of memes (sorry). From his position as a neuroscientist, he claims that the shared architecture of human brains allows certain memes to take hold and synchronize individuals at the neuronal level, accounting for everything from the Enlightenment to the willingness of British and German soldiers to charge out of their trenches to certain death during the Battle of the Somme. I have no doubt that this true, but I also think that some sort of neuronal link was implicit in Dawkin's original formulation (where was the meme supposed to gestate, in a box?), and the concept of brainets doesn't seem to add much to Nicolelis' overall argument.

This is disappointing, because earlier in the book, Nicolelis hints at a much more original definition of brainets. This stems from his experiments with actual brain-to-brain coupling in test animals, so that, for example, one rhesus monkey (with titanium pins in its brain) can help another rhesus monkey (also with titanium pins in its brain) navigate an improvised wheelchair toward a pre-established goal, so both can reap the reward of a tasty fruit drink. If this sounds creepy, that's because it is. To paraphrase another critic of Nicolelis' work, it's like the opening scene of a Paul Verhoeven movie the studio decided not to release.

Why does Nicolelis lay so much stress on memes (I mean, brainets?) Because he's building toward his main theme, that human brains, both individually and collectively, determine the reality of our shared existence. First, though, we have to deal with the small issue of life, the universe, and everything.

**[Question #3:** How do you feel, personally, about inserting titanium pins into the brains of live rhesus monkeys? I have to admit, I'm a little freaked out about it.

## What is Mind? No Matter. What is Matter? Never Mind

Remember that analogy I quoted above, where you have to take into account a pool player's personal gravitational field in order to calculate the knock-on effects of his opening break? That is how I feel about books like *The True Creator of Everything*. The author starts with one-reasonable sounding concept, then adds another reasonable-sounding concept on top of it, and by the end of the book his inexorable chain of reasonable-sounding arguments lands you somewhere that...doesn't seem quite so reasonable. But the scenery is interesting, so at least let's talk about it.

After an interesting discussion about how the introduction of clocks in the Middle Ages permanently altered human beings' relationship with and perception of time, and how Newton

and Einstein's scientific discoveries permanently altered human beings' relationship to and perception of both time and space, Nicolelis offers this:

*The very primitive concepts of space and time are also mental abstractions created by the human mind in order to reduce the dimensionality of complex potential information obtained from the outside world. Furthermore, I propose that as basic mental abstractions, time and space emerge as a result of natural selection...as a way of enhancing our evolutionary fitness. By filling the human universe with a continuous scaffold made of time and space, our brains enhance our chances to survive the contingencies imposed by the environment in which we have been immersed since the origin of our species.*

This is not the first time I've heard this idea, but it may be the first time I've heard it expressed by a reputable scientist. Nicolelis adduces, as evidence for our continuing "creation" of space and time, the ability of the human mind to fill in information that isn't actually there: perceiving a solid triangle in an optical illusion, filling in missing gaps in conversation, basically making sense of a world where solid information is at a premium. Nicolelis goes so far as to speculate that hallucinogens work not by impairing brain tissues that ordinarily process normal reality, but by reducing the ability of these tissues to manufacture the illusion of space and time out of the chaos that lies outside human perception.

But wait, it gets even weirder. Over and above organic computers, brainets, and the manufacturing of space and time, the most thought-provoking idea in *The True Creator of Everything* (in my opinion) concerns the true nature of mathematics. Here, Nicolelis relies heavily on *Where Mathematics Comes From: How the Embedded Mind Brings Mathematics into Being*, by George Lakoff and Rafael Nunez. This book, which I recommend, makes a strong case that human mathematics is based on metaphorical aspects of everyday human existence—we talk about objects residing *within* a set, being located a certain distance *along* a number line, having certain geometrical properties based on our experience of three dimensions. There's a strong case to be made that mathematics is not platonic in nature, but is simply another human creation, like music or architecture. As Nicolelis puts it, "the whole body of accumulated mathematical knowledge could be seen as another type of emergent property produced by a human brainet, dispersed in time and space, over the entire history of humankind."

Nicolelis proposes—I'm not saying I agree with him, I'm just saying he proposes—that any intelligent species arising in any universe would necessarily have a different evolutionary history and brain structure than us human beings, and, to mirror his sentence in my own words, "the whole body of accumulated mathematical knowledge could be seen as another type of emergent property produced by a Phlegmorphian brainet, dispersed in time and space, over the entire history of Phlegmorphkind."

The implications of this are stunning, provided (with a big P) you accept Nicolelis' arguments. Why should we have any expectation that, with their different brain structure and shared evolutionary history, Phlegmorphians would perceive time and space the same way we do, or have the same kind of mathematics? What Nicolelis seems to be proposing is that not only would it be *\*very difficult\** for humans to communicate with another intelligent life form in the universe (think of those octopus beings in the movie *Arrival*), but it might actually be *\*logically impossible\** to do so.

**[Question #4:** Is there any ontological/epistemological/phenomenological etc. content to the notion that another intelligent species in the universe could evolve a type of mathematics or cosmology that is completely incommensurable with our own? This sounds to me like trying to

run away from your own shadow—I literally can't conceive that it's possible, but does that mean it's impossible?]

This leads me to a final big idea, which is not in Nicoletti's book, but which is mine because I thought of it and it is my theory and no one else's. What if the solution to the Fermi Paradox is that intelligent observers actually do create their own reality, as Nicoletti says, and this reality is incommensurable with the reality of any other intelligent observers? Then, by definition, our universe could only be inhabited by one intelligent species—us, and our big brains.

# The WEIRDest People in the World by Joseph Henrich

[RL]

*The cultural evolution of psychology is the dark matter that flows behind the scenes throughout history.*

Joseph Henrich is the world's leading scholar of cultural evolution. He released a brilliant book last week, *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*.

The book starts from this question: Why are Western folks psychologically different than other people around the world?

This Western peculiarity was found after scientists realized that many of their research subjects were homogenous: Western, Educated, Industrialized, Rich, and Developed. Henrich calls these folks "WEIRD". Psychologically speaking, WEIRD folks are especially individualistic and analytical. In contrast, non-WEIRD folks (from Asia or the Amazon) are more collectivist and holistic.

Henrich asks the question: **Why are WEIRD folks so peculiar? And how has this WEIRDness made the West so prosperous?**

He tackles this question by exploring the co-evolutionary process between genetics, biology, psychology, norms, and institutions. Henrich looks at how these processes co-evolve through time: from apes, to early homo sapiens, kin-based structures in the Agrarian Age, the Catholic church in the Middle Ages, the Protestant Reformation, the Industrial Revolution, and finally to the modern institutions of our modern era.

Let's start by giving an overview of Henrich's model—how genetics, biology, psychology, norms, and institutions co-evolve.

## I. Systemic Evolution

Systemic evolution (my term, not Henrich's) is the co-evolutionary process outlined above—how biology co-evolves with culture. It is the unique lens that cultural evolution provides as an academic field. For a simple example of this lens, let's consider how our *genes* have co-evolved with our *culture*.

*Why do folks in Nordic countries have blue eyes?* First, early homo sapiens had to survive in the cold north. To do so, we developed the technology to farm cereal. This allowed us to move north and still eat. But the north doesn't have much sun, which means we get less Vitamin D and don't need to worry as much about UV skin damage. So then we started to produce less melanin and our skin got whiter. As part of this, one of our melanin-producing genes had a mutation which made it produce less melanin. This made our skin whiter. And coincidentally, that "skin color" gene was close to our "eye color" gene, and so we got blue eyes!

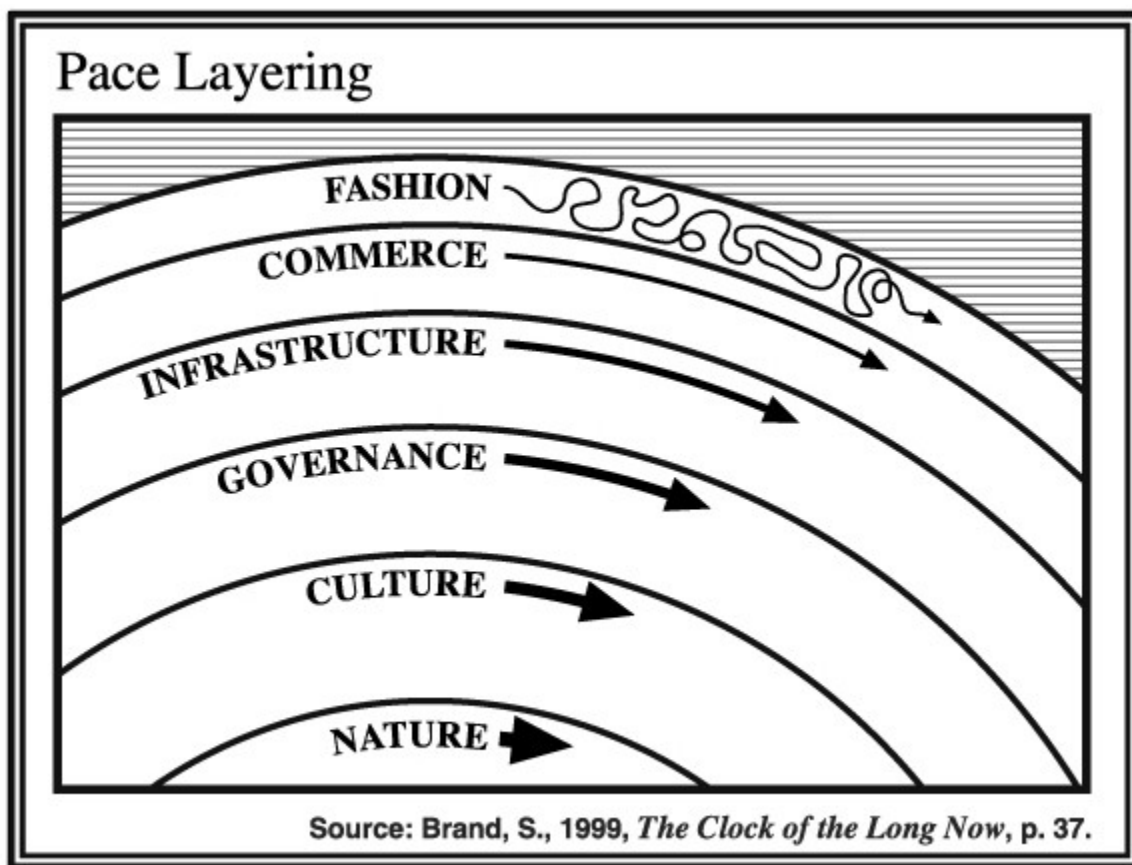
Before humanity developed farming technology, all of us had brown eyes. Then we developed an adaptation at the **cultural level** (farming), which changed us at the **genetic level** (blue eyes). This example of culture-gene co-evolution points to the essence of systemic evolution—that our biology evolves as our society evolves.

There are five levels of systemic evolution: genetics, ontogeny, psychology, norms, and institutions.

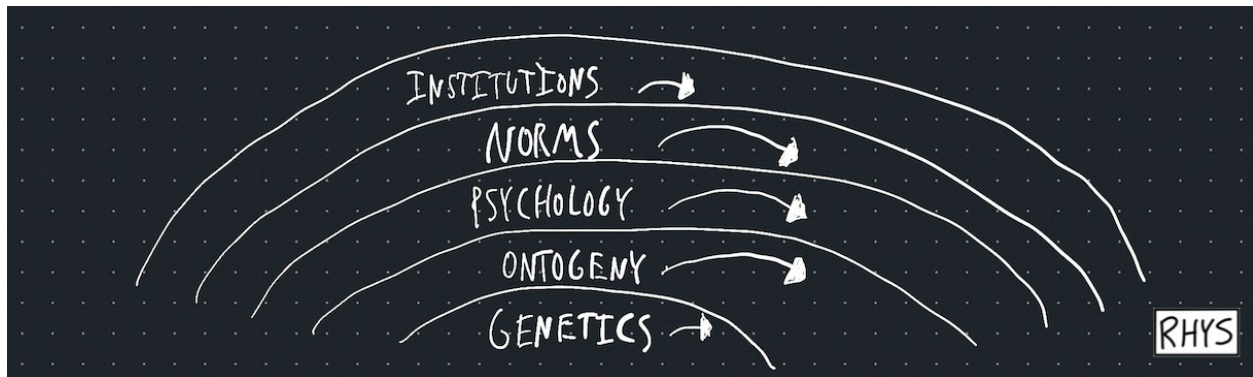
- **Genetics:** Genetic evolution happens slowly on generational time-scales of thousands of years. "Fitness" is determined by what genes reproduce the most.
- **Ontogeny:** Ontogeny refers to the physical development of an organism *during its own lifetime* while genetics (or phylogeny) refers to how the organisms have evolved across lifetimes. As an example, taxi drivers have highly developed hippocampi, the part of the brain responsible for navigation.
- **Psychology:** Psychology is how our brains process information. As opposed to the levels above, which are purely biological, psychology is more culturally-focused. As an example, WEIRD societies are individualistic, while non-WEIRD societies are collectivist.
- **Norms:** Norms are what happens when we project our inner psychology out into the world. It is the praise we receive for jumpstarting a neighbor's car. Or the shame when someone steals a bike. These norms can become moderately calcified into habitual actions like shaking hands to greet each other.

- **Institutions:** Institutions are the formalization of norms into even more calcified structures like laws, markets, corporations, and nation-states. These institutions are usually written down and change more slowly.

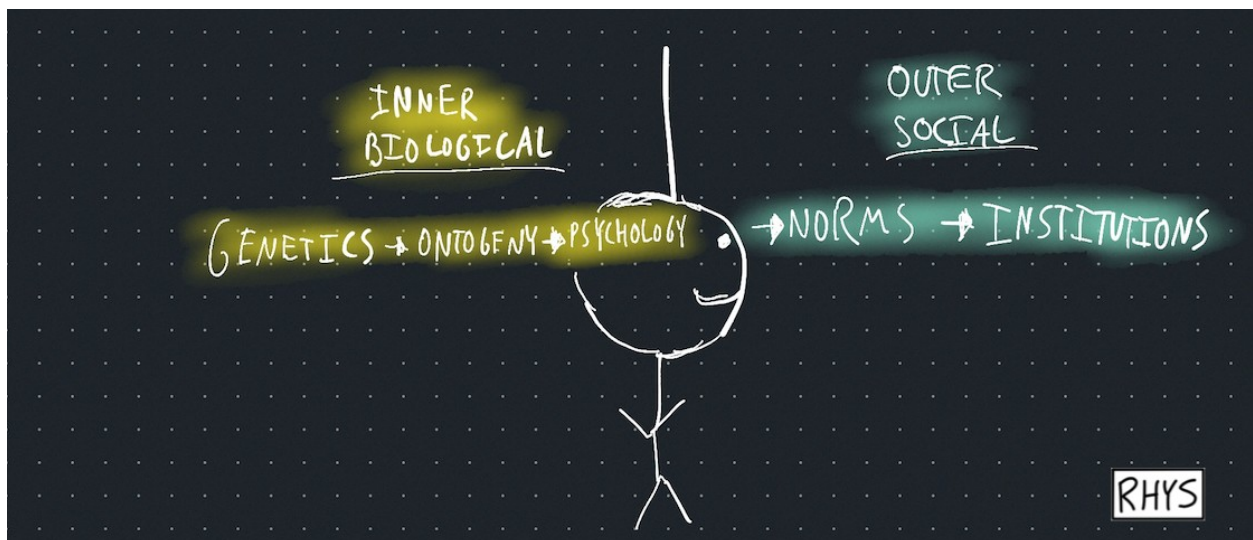
I like to think of these levels like Stewart Brand's pace layers—they move at different time scales. Here's Brand's pace layering:



And here's Systemic Evolution, which moves in a slow-fast-slow pattern:



In addition, System Evolution moves from the inner biological world, to the outer social world:



How can we understand the evolution of these layers? To do this, we need to understand the how higher cultural layers in the evolutionary stack "fit" onto lower down biological anchors.

## II. Biological-Cultural Fit

**Biological-cultural fit** is the extent to which cultural adaptations fit onto our biological anchors. For example, we can't make biological ants believe in a Sun God. Ants' simple biology isn't a good anchor a complex culture.

For another example, let's look at how our *cultural* marriage norms were built on our *biological* pair-bonding instincts. As homo sapiens moved away from apes, we developed pair-bonding instincts to have a male-female pair watch over their children as their brains developed. From that *biological* instinct (pair-bonding), we developed the *cultural* technology to create marriage norms like wedding vows. But if we didn't have



our pair-bonding anchor, those marriage norms wouldn't have stuck. Imagine trying to get sex-happy polygynous bonobos to settle down with one partner for life!

We can only jump so much in our evolution at any given time. This leads to a certain kind of "path dependence". Homo sapiens in 2020 are who we are because of homo sapiens in 10,000 BCE. And homo sapiens in 3000 can only follow from homo sapiens in 2020.

In technology, we'll often call this the "law of the adjacent possible." The iPhone wasn't possible in 1930s before computers. We needed information theory, transistors, digital cameras, and GPS before we could build the iPhone.

As another technological metaphor, think of Product-Market Fit (PMFit). PMFit says that a product will only be successful if it meets the needs of the market. Similarly, a given cultural innovation is only possible if it fits our natural biological anchor. We could call this Biological-Cultural Fit, or BCFit.



Now that we've understood Systemic Evolution and BCFit, we can trace these over time to understand how the West got so WEIRD.

### **III. Apes (Pre-200,000 BCE)**

We started as apes. Now we're apes who wear clothes!

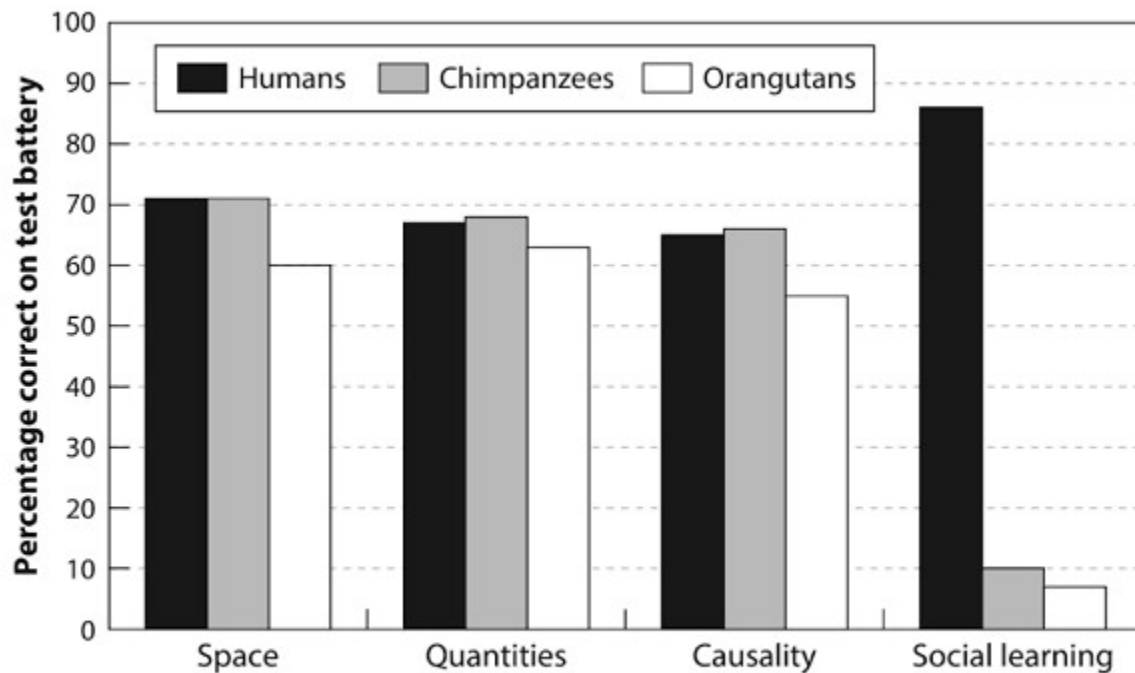
As apes, we didn't have the solidified norms and institutions that we've developed today. Those are inventions of homo sapiens.

However, we did have two important biological instincts which provide the anchor for future cultural evolution:

- **Polygynous Mating:** All ape species besides homo sapiens engage in polygynous mating. All future mating/marriage norms need to be built on our initial anchor of polygyny.
- **Kin Altruism:** Other ape species engage in kin altruism, where individuals cooperate with close kin in order to increase genetic fitness of close relatives. For example, I may give some extra food to my hungry brother because it increases the chances of his reproduction (and doesn't hurt me too much). Many of our kin-based institutions (like patrilineal clans) are built on the biological anchor of kin altruism.

#### **IV. Early Homo Sapiens (200,000 BCE - 10,000 BCE)**

Homo sapiens began to branch off from our ape ancestors. Many of the evolutions around this time were based on social learning and homo sapiens big brains. In fact, our ability to do social learning is what differentiates us from our ape ancestors. In the graph below, we can see how human toddlers compare to chimpanzees and orangutans on a series of cognitive tests. As you can see, humans are roughly equal to apes on tests of physical space, quantity, and causality. But we're much better than apes at social learning.



**Figure 2.2.** Average performance on four sets of cognitive tests with chimpanzees, orangutans, and toddlers.

Homo sapiens brains are optimized for social learning. Instead of being "smart" when we come out of the womb, we start "dumb" but then pick up cultural adaptations extremely quickly. It's what makes us so flexible. Thousands of years ago, we could learn animal tracking from our community. Now, we learn the base-10 counting system.

Genetically, our bodies adapted for new big brains with things like wider hips for females to birth big brains. Culturally, we adapted by developing **pair-bonding instincts**. These instincts encourage the male-female sexual pair to bond and parent the child during their formative early years.

From a BCFit perspective, pair-bonding instincts were built on the polygynous anchor described above. This means that our (monogamish) pair bonds swim upstream of our polygynous nature. And as we'll soon see, these pair-bonding instincts set the anchor for later marriage norms.

## V. Kin-Based Structures in Agrarian Age (10,000 BCE - 500 CE)

Around 10,000 BCE, humans learned how to domesticate animals and plants for farming. This technological evolution led us to larger, more complex societies with proto-institutions like states.

But these institutions were *not* like the Western institutions we have today, which are built on impersonal trust of strangers. Instead, we built institutions based on our kin instincts. These kin-based institutions had stronger Biological-Cultural Fit. Almost all of these kin-based institutions were based on some form of unilineal or patrilineal descent.

### **Patrilineal Descent**

**Patrilineal descent** is the cultural adaptation that we should organize our kin structures around our male lineage. In Western societies, children take the father's last name. Or last names themselves will show your lineage: Johnson, Wilson, Anderson, etc.

Why did patrilineal descent show up? Remember, it was *evolved*, not *invented*. It's not like some old rich white guy decreed "one must track descent through the father." As it evolved, it needed to *outcompete* the exist system of bilineal descent (tracking both mother and father lines).

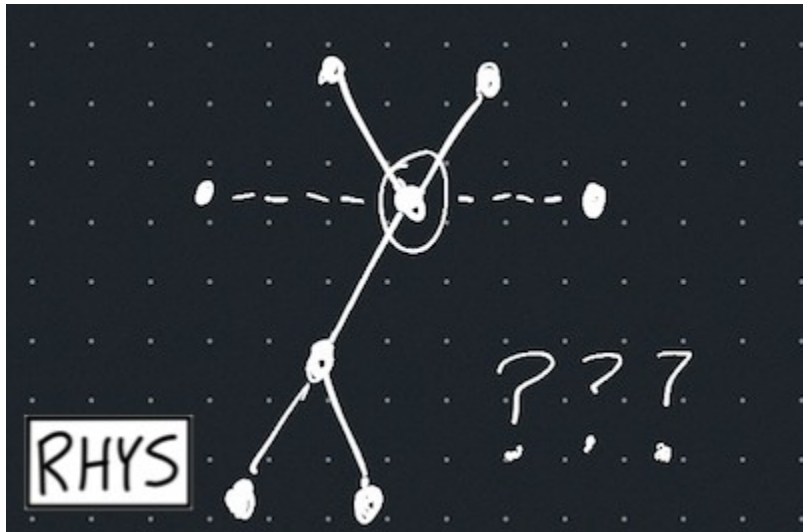
Patrilineal descent outcompeted bilineal descent because it mitigated conflicts of intra-family interest while also providing clear lines of authority. To see the difficulties of bilineal descent, Henrich gives the example of a hunting party:

*To see bilineal conflicts, suppose we start with a father who is putting together a defensive party of 10 men to drive some interlopers off their community's land. The father, Kerry, starts by drafting his two adult sons. This is a nice trio, evolutionarily speaking, since not only are all three closely related but they are also equally related—fathers and sons are genetically related at the same distance as brothers. This parity minimizes conflicts of interest within the trio.*

*Now, Kerry also recruits his older brother's two sons, and their sons, who are just old enough to tag along. Still short three men, Kerry recruits his wife's brother, Chuck, and his two sons.*

*As you can see, this is a mess of potential conflicts with several possible cleavages. What if Chuck faces a choice between saving one of his own sons in the melee or Kerry's two nephews? What if Kerry's nephew gets one of Chuck's sons killed?*

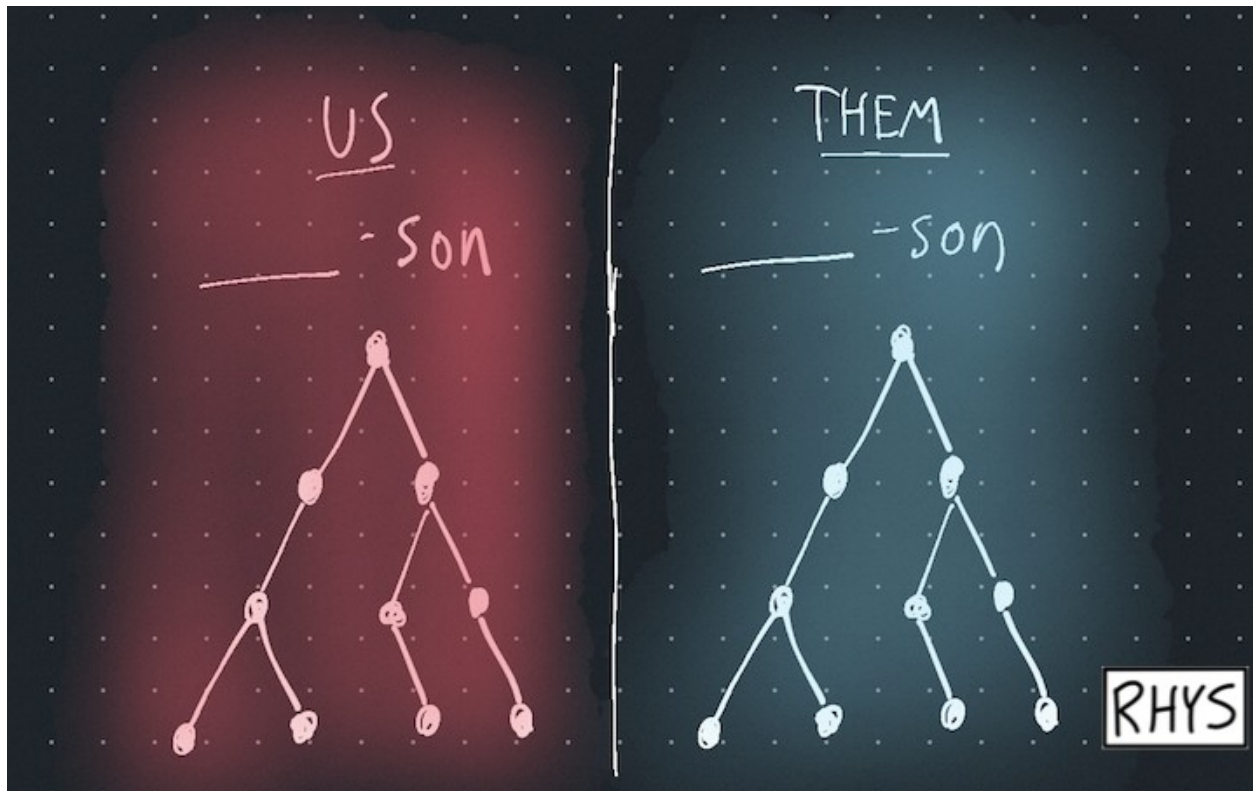
This is why bilineal descent is confusing. Who is on your team? And who do you take instructions from?



Instead, we can use patrilineal descent to reduce the conflicts of interest. Henrich writes:

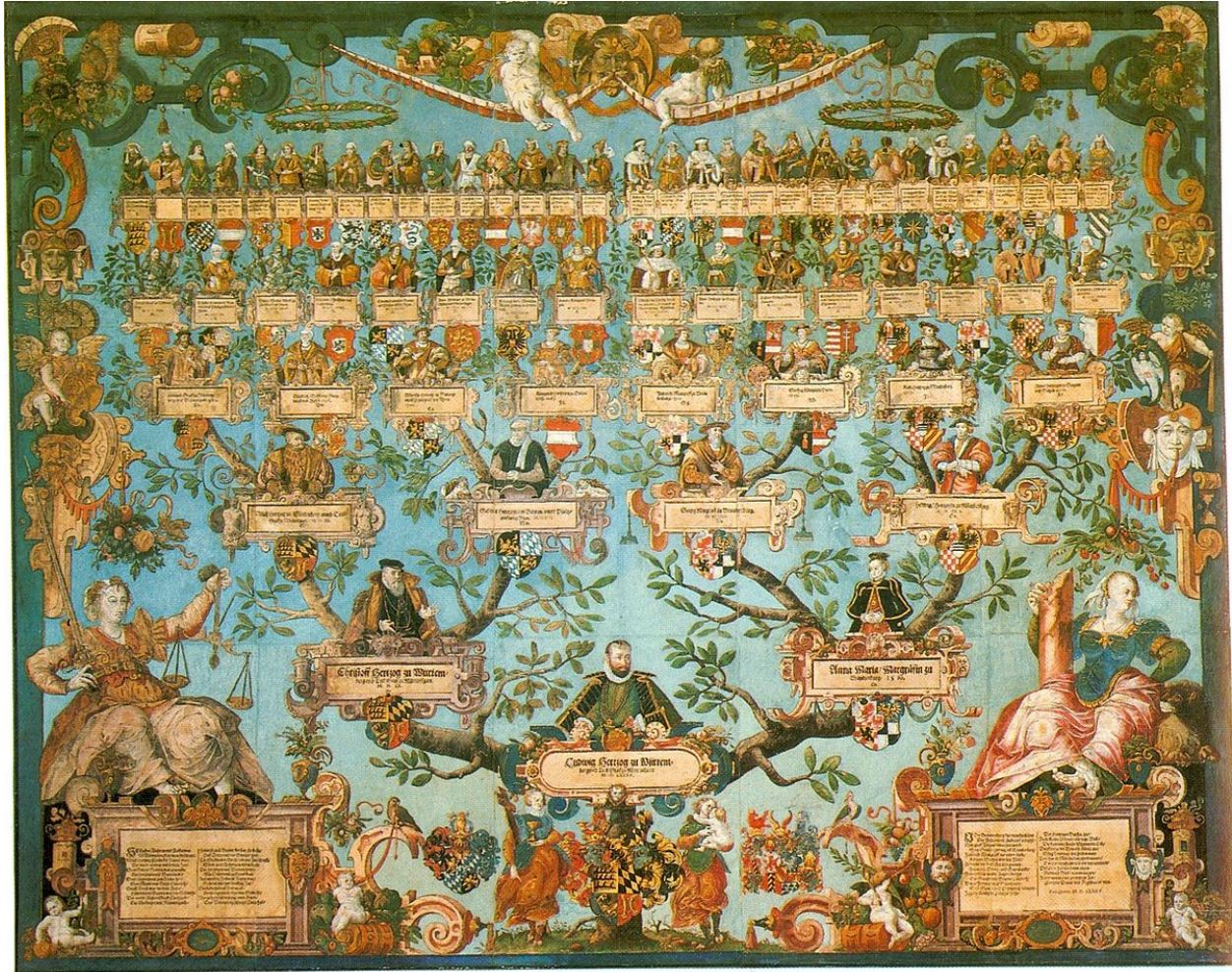
*To mitigate such conflicts, clans elevate one side of a person's genealogy over the other and shift the focus of kinship reckoning from one centered on each individual to one centered on a shared ancestor. Thus everyone from the same generation is **equally related** to a shared ancestor, and everyone has the **same set** of relatives. This notion is amplified in how these societies label and refer to their relatives in their kinship terminologies. In patrilineal clans, for example, your father's brother is often also called "father".*

Patrilineal descent simplifies kin structures by turning them into an us vs. them dynamic. It's \_\_\_\_-son vs. \_\_\_\_-son (Anderson vs. Johnson).



Patrilineal descent birthed a variety of social norms like patrilocal residence (which builds bonds among a father's children), equal-stake inheritance (so all children have aligned financial incentives), incest taboos (which decrease sexual competition within the clan), and arranged marriages (to create a network of alliances). In addition, humans created **segmentary lineages**—genealogical trees with deep rituals to remember the full clan.





By 500 CE, patrilineal descent had created a world full of strong kin-based clans. These clans had built kin-based norms and institutions on top of our biological tendency towards kin altruism and pair bonding.

However, our modern 2020 society is full of non-kin-based institutions like companies, governments, universities, and organized religion. How did we move away from kin-based institutions? The Catholic Church.

## VI. Catholic Church in the Middle Ages (500 CE - 1500 CE)

During the Middle Ages, the Catholic Church both: a) broke down existing **kin-based** institutions and b) built up new **non-kin-based** institutions. Let's look at each.

### VI.A How the Church's Marriage and Family Program (MFP) Broke Kin-Based Institutions

When it formed, Christianity was a hip new kind of religion—one with a moralizing god. Previous religions didn't care much about your personal actions. But Christianity (and

Islam, among others) had found a cultural adaptation. By saying "if you do good, God will get you into heaven", these new religions were able to increase impersonal trust between strangers, which allowed larger, more complex societies to form.

As these religions began to spread, they competed with existing kin-based institutions. The Catholic Church competed with kin-based clans by: a) Breaking kin lineages and b) Redirecting financial inheritance from kin to the church.

The Church broke kin lineages by exploiting their Achilles' heel—they must produce heirs every generation. Henrich writes:

*A single generation without heirs can mean the end of a venerable lineage. Mathematically, lineages with a few dozen, or even a few hundred, people will eventually fail to produce an adult of the "right" sex. This means that all lineages will eventually find themselves without any members of the inheriting sex. Because of this, cultural evolution has devised various strategies of heirship that involve adoption, polygamy, and remarriage.*

The Church exploited this weakness by breaking kin lineages with new rules that constrained adoption, polygamy, and remarriage (their "Marriage and Family Program").

In addition, the Church took money away from kin-based lineages by encouraging inheritance donations to the Church. They even provided a powerful carrot: if you donated to the Church, you were more likely to go to Heaven. And, instead of requiring this during life, you could do it at the end:

*Rich people could bequest some or all of their wealth to the poor at the time of their death. This allowed the wealthy to stay rich all their lives, but to still thread the proverbial needle, by giving generously to the poor at their death.*

This was a massive moneymaker for the Church. Bequests of land were the most common:

*By 900 CE, the Church owned about a third of the cultivated land in western Europe. By the Protestant Reformation in the 16th century, the Church owned half of Germany, and between one-quarter and one-third of England.*

Over the course of centuries, the Church slowly broke down kin-based lineages by: a) Destroying the lineages themselves and b) Redirecting their sources of income.



## VI.B The Creation of Non-Kin-Based Voluntary Organizations

The Church dismantled kin-based tribes that had previously provided humans with interdependence networks that met our needs of safety and connection. But we still needed to meet those needs. What would take their place?

At this time, a wide variety of non-kin-based voluntary associations began to form in cities: companies, churches, guilds, unions, political parties, and universities. These looked like similar organizations in the East, but were actually built on a new proto-WEIRD foundation. Henrich writes:

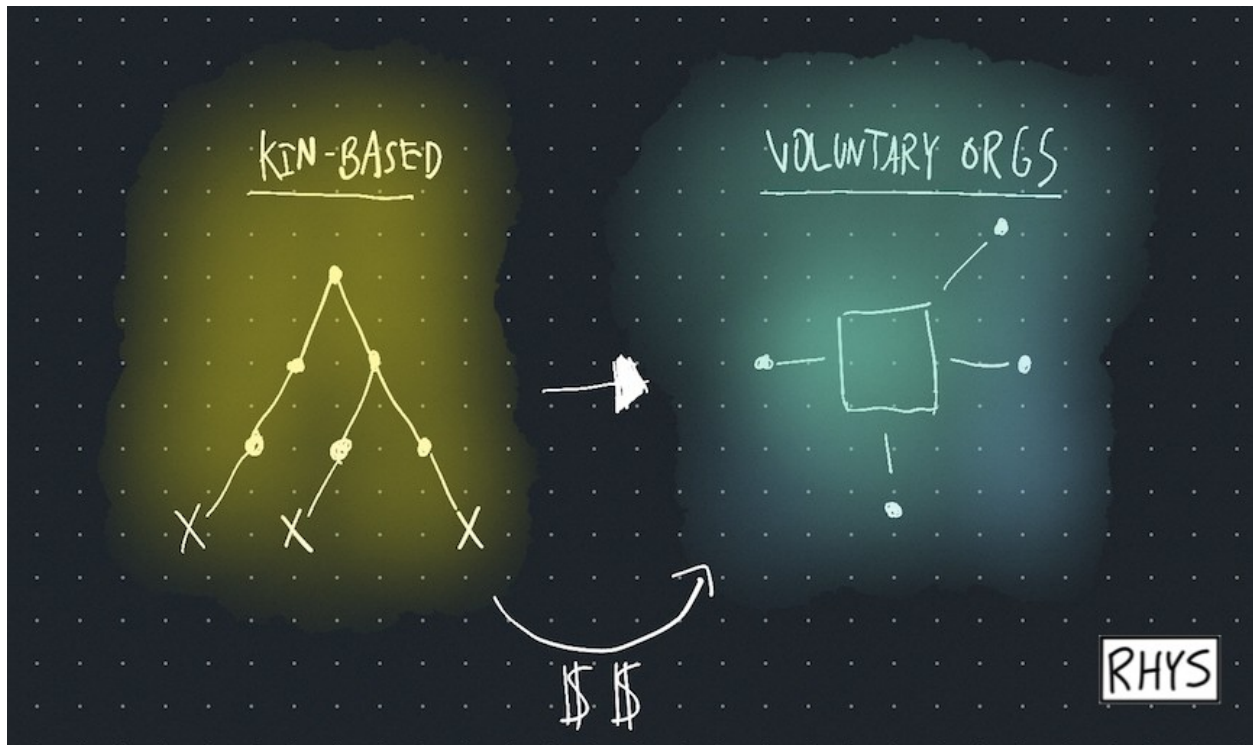
*While the urban centers of 11th-century Europe may have superficially looked like puny versions of those in China or the Islamic world, they were actually a newly emerging form of social and political organization, ultimately rooted in, and arising from, a different cultural psychology and family organization. Smaller families with greater residential and relational mobility would have nurtured greater psychological individualism, more analytic thinking, less devotion to tradition, stronger desires to expand one's social network, and greater motivations for equality over relational loyalty.*

From 800-1800 the urban population in Europe increased from 5% to 20%. As these cities formed, they created charters and laws to regulate urban behavior. Popular charters (like Magdeburg Law) were copied and remixed throughout Europe.

In these newly booming cities, colleges began to form. The first modern university started in Europe in 1000. 500 years later, there were 50 universities across Europe.

Markets spread as well. Cities applied for grants to hold violence-free markets in their jurisdictions. At the same time, these cities began to attract merchants, traders, lawyers, and other professionals to engage in commerce. They developed *lex mercatoria*, a set of norms and laws around impersonal exchange.

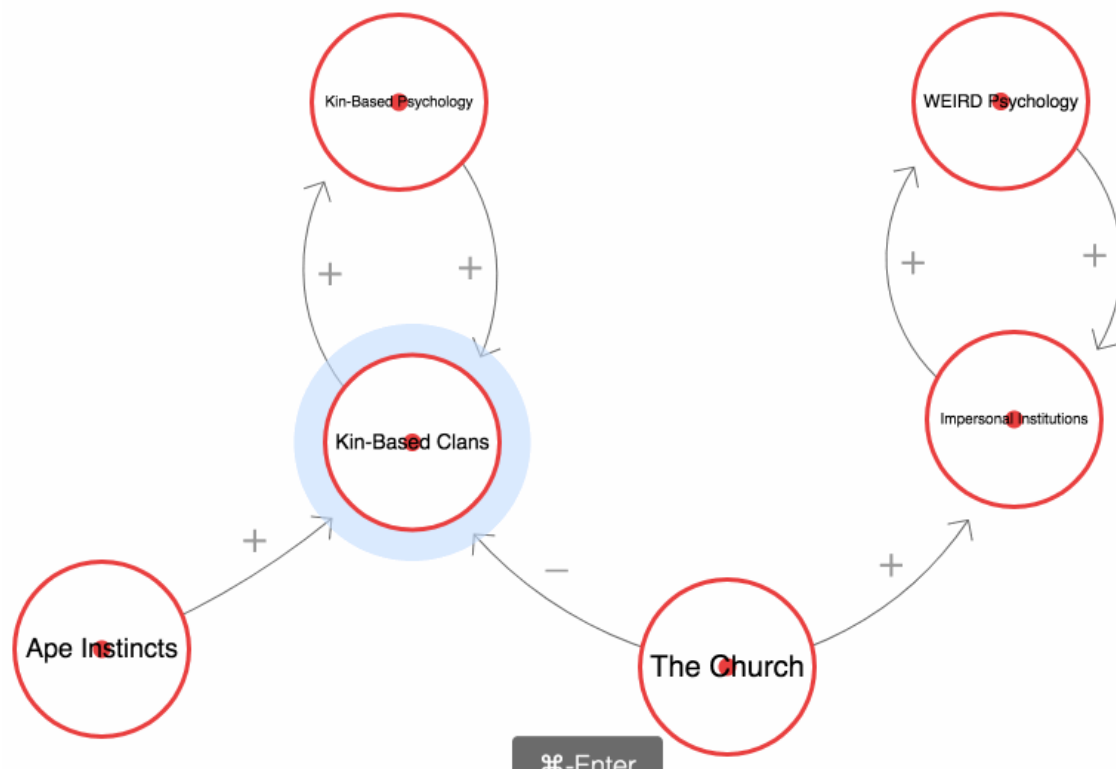
These voluntary organizations co-evolved with a proto-WEIRD psychology that elevated the individual over the collective, singular objects over holistic relationships, and self-focused guilt over community-given shame. In our previous kin-based culture, it was incredibly important to think of oneself in relationship to the personal contexts around us. But as voluntary associations began to form, it was increasingly important for us to provide a clear individual outward face for others to interact with, like an API.



These voluntary organizations were possible because of *both* the destruction of old kin-based culture *and* the incubation of an individualistic proto-WEIRD psychology that made individuals more likely to adopt impersonal associations.

In addition, these voluntary organizations (and their accompanying proto-WEIRD psychology) provided the anchor upon which the Protestant Reformation, Industrial Revolution, and modern nation-states would form.

This is all shown in the feedback loop below. Ape instincts led to the co-evolution of kin-based clans and psychology. Then The Church broke that and led to the co-evolution of WEIRD institutions and psychology.



## VII. Protestant Reformation and the Industrial Revolution (1500-2000)

Protestantism sacralized the psychological complex that had been percolating in Europe during the centuries leading up to the Reformation.

In other words, the Protestant Reformation took our proto-WEIRD individualism and codified it into religion:

Embedded deep in Protestantism is the notion that individuals should develop a personal relationship with God. To accomplish this, both men and women needed to read and interpret the Bible for themselves, and not rely primarily on the authority of supposed experts, priests, or institutional authorities like the Church. This principle is known as sola scriptura.

As Max Weber argues in *The Protestant Ethic and the Spirit of Capitalism*, this Protestant ethic led directly to the Industrial Revolution and modern capitalism.

But was this the only explanation for why Europe was the center of the Industrial Revolution? No. Henrich writes:

*Proposed explanations for “Why Europe?” emphasize the development of representative governments, the rise of impersonal commerce, the discovery of the Americas, the availability of English coal, the length of European coastlines, the brilliance of Enlightenment thinkers, the intensity of European warfare, the price of British labor, and the development of a culture of science.*

Henrich believes it was our proto-WEIRD psychology that underpinned each of these explanations:

*I suspect that all of these factors may have played some role, even if minor in some cases; but, what’s missing is an understanding of the psychological differences that began developing in some European populations in the wake of the Church’s dissolution of Europe’s kin-based institutions.*

*The cultural evolution of psychology is the dark matter that flows behind the scenes throughout history.*

Let's look at a few manifestations of this psychological WEIRD dark matter: a market mindset, individual rights, and scientific practices.

## **VII.A Doux Commerce as Domesticated Intergroup Competition**

Capitalism rose with a market psychology that Henrich touches on. (And, imo, is most clearly explained in Albert Hirschman's *The Passions and the Interests: Political Arguments for Capitalism Before Its Triumph*.) The idea is one of domesticated intergroup competition. Instead of killing each other in competition, we just beat each other in the market.

This mindset (of gentle commerce, or *doux commerce*) was spearheaded by various thinkers at the time:

*Commerce is a cure for the most destructive prejudices; for it is almost a general rule that wherever manners are gentle there is commerce; and wherever there is commerce, manners are gentle.*

—Montesquieu (1749)

*Commerce is a pacific system, operating to cordialise mankind, by rendering Nations, as well as individuals, useful to each other...The invention of commerce...is the greatest approach toward universal civilization that has yet been made by any means not immediately flowing from moral principles.*

—Thomas Paine (1792)

## VII.B Individual Rights and Democracy

In addition, our proto-WEIRD individualism also led to the formation of modern law and especially the concept of **individual rights**. Henrich again:

*The Declaration of Independence asserts, “We hold these truths to be self-evident, that all men are created equal, that they are endowed by their creator with certain unalienable rights, that among these are life, liberty, and the pursuit of happiness.”*

*If the idea that people are endowed with such abstract properties makes sense to you, then you are at least a little WEIRD.*

*By contrast, from the perspective of most human communities, the notion that each person has inherent rights or privileges disconnected from their social relationships or heritage is not self-evident.*

*And from a scientific perspective, no “rights” have yet been detected hiding in our DNA or elsewhere. This idea sells because it appeals to a particular cultural psychology.*

Individual rights gave individuals power they had never had before. In addition, democracy itself began to empower the general populace. Before democracy, only gods or lineage gave one legitimacy. Now, the people themselves were seen as a source of legitimacy. (Crazy!)

## VII.C Individualistic Science and the Collective Brain

With the rise of universities, European science was beginning to blossom. These new European scientists were a bit different though—they had proto-WEIRD individualism. Instead of respecting their scientific elders, they now actively pushed back against them. This led to paradigm shifts like Copernicus' discovery of heliocentrism in 1543.

In fact, the very notion of "discovery" was discovered at this time. The word itself first appears in European languages around this time: Portuguese in 1484, Italian in 1504, etc.

In addition, scientists began to associate discoveries with individuals. Henrich writes:

*Our commonsensical inclination to associate inventions with their inventors has been historically and cross-culturally rare. This shift has been marked by the growth of eponymy in the naming of new lands (“America”), scientific laws (“Boyle’s Law”), ways*

*of thinking ("Newtonian"), anatomical parts ("fallopian tubes"), and much more. After about 1600, Europeans even began to relabel ancient insights and inventions based on their purported founders or discoverers. "Pythagoras's theorem," for example, had been called the "Dulcarnon" (a word derived from an Arabic phrase for "two-horned," which described Pythagoras's accompanying diagram)."*

Ideas themselves could now be stolen:

*"Marking this in English, words for "plagiarism" first began to spread in the 16th century, following the introduction in 1598 of the word "plagiary," which derives from the Latin word for kidnapping."*

With the rise of the printing press and knowledge societies, these new individualistic scientists began to operate as a collective brain. Henrich again:

*The Cistercian Order, in particular, built a sprawling network of monastery-factories that deployed the latest techniques for grinding wheat, casting iron, tanning hides, fulling cloth, and cultivating grapes. At mandatory annual meetings, hundreds of Cistercian abbots shared their best technical, industrial, and agricultural practices with the entire order. This essentially threaded Europe's collective brain with Cistercian nerves, pulsing the latest technical advances out to even the most remote monasteries."*

---

These examples—individual rights, *doux commerce*, and an individual collective brain—are all manifestations of this new individualistic WEIRD psychology. Can we get closer though? To measure WEIRDness itself?

The most direct way to measure this psychology is through the BIG-5 personality test, which is a scientifically-backed set of 5 traits that all individuals can be measured by. You can remember them with the mnemonic OCEAN: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism.

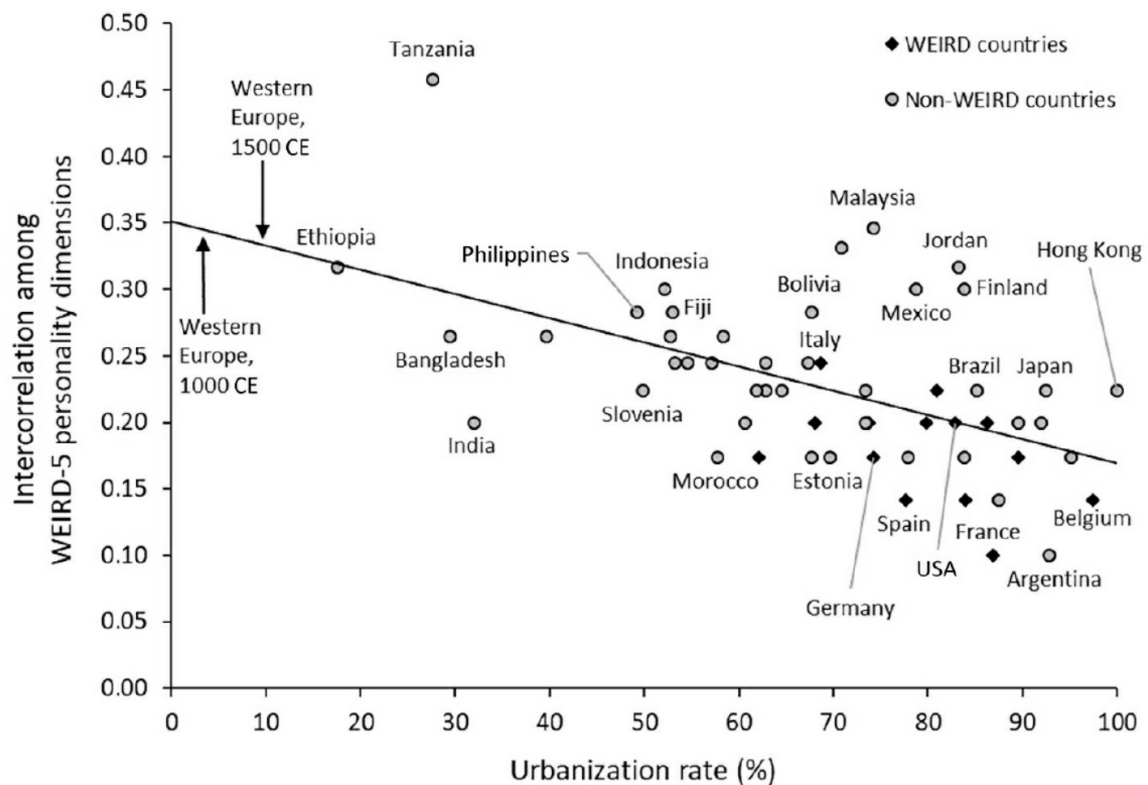
So, we'd expect WEIRD folks to look different than non-WEIRD folks, right? Like maybe Western folks are more extraverted?

Wrong! In fact, non-WEIRD folks don't even register on the same personality test—they have a completely different set of personality dimensions. For example, some anthropologists studied the remote Tsimané people. Their society is built differently than Western society (i.e. no impersonal institutions). And therefore, the personality types that one can "be" in their society are different. Instead, their personality types mostly group along two clusters: pro-sociality and industriousness. In rural Tsimané society,

there's no evolutionary niche for 5-dimensions of personality to emerge. While in a more urban culture, there's room for occupational diversity.

We can see this statistically by looking at the intercorrelation among BIG-5 traits. In theory, the traits should be orthogonal to each other. For example, in the USA, there's only a 0.10 correlation between traits. But in a non-WEIRD country like Tanzania, there's nearly a 0.50 correlation between traits!

The graph below (from Henrich) shows how urbanization rate correlates with BIG-5 intercorrelation.

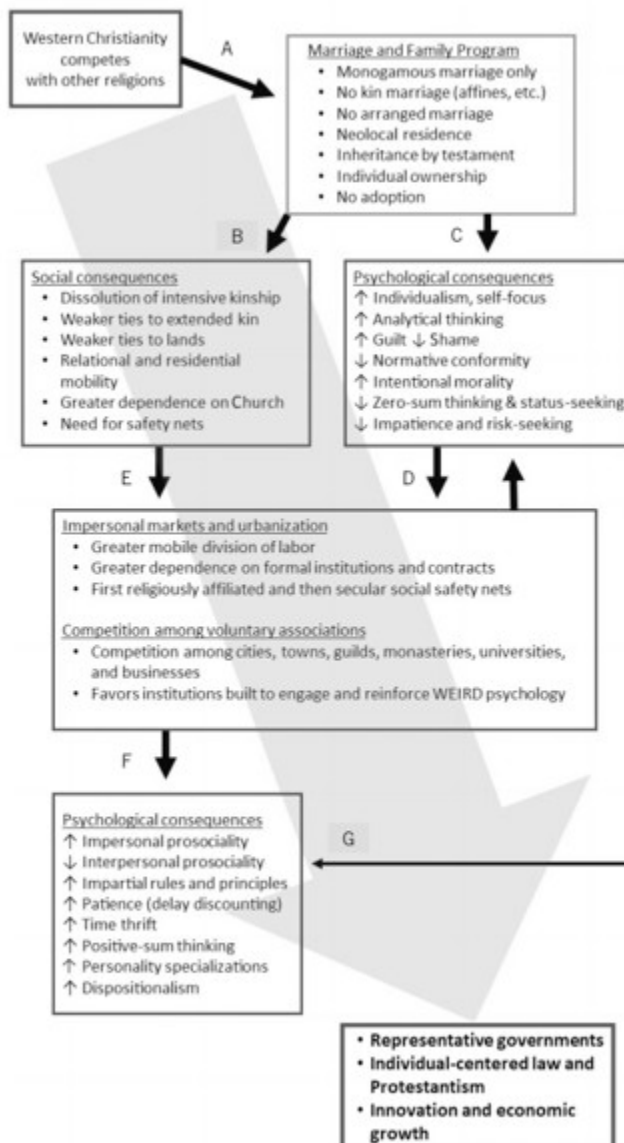


This is all to say: BIG-5 personality traits are just a manifestation of WEIRD psychology and parallel co-evolution of WEIRD impersonal institutions. It's not the BIG-5, it's the WEIRD-5.

## VIII. Conclusion

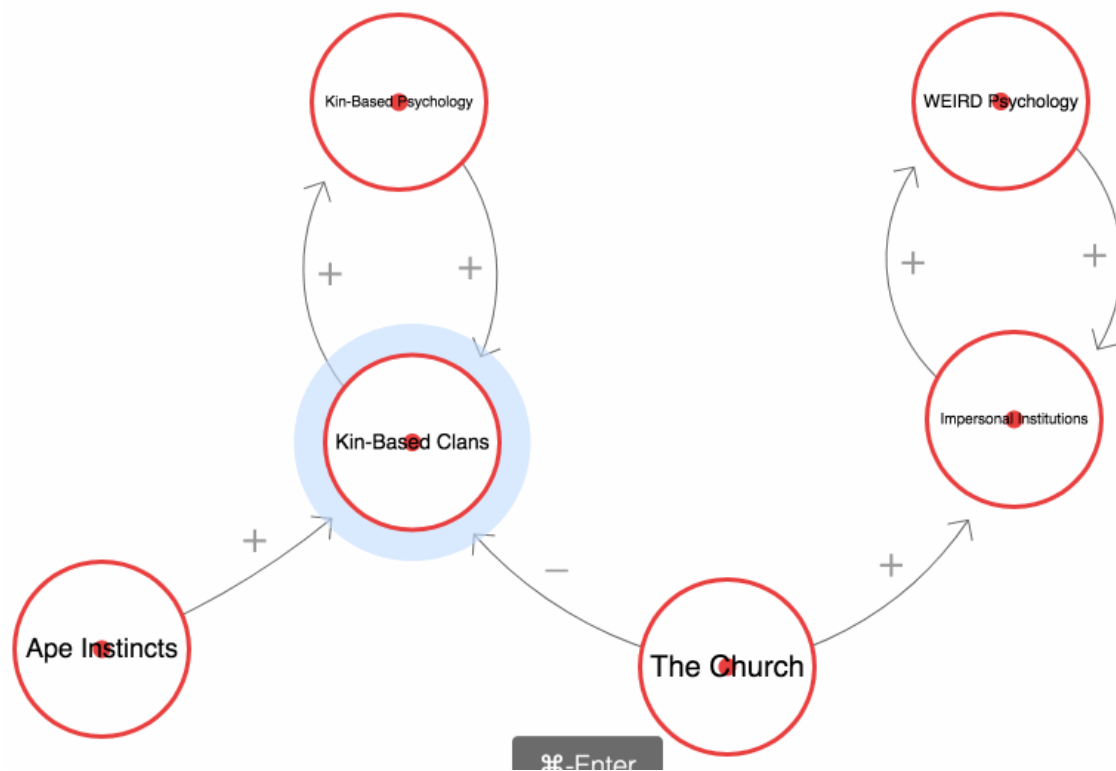
In today's article, we looked at why the West is so WEIRDly individualistic: The Church broke down kin-based structures which created space for impersonal institutions that

highlighted the individual. These new institutions became embedded in people as a WEIRD individualistic psychology. (Which then led to even more impersonal institutions.) Here is how Henrich explains the whole trajectory:



But I'd represent it as this Loopy. Ape instincts led to the co-evolution of kin-based clans and psychology. Then The Church broke that and led to the co-evolution of WEIRD institutions and psychology.





In the next part of this series, I'll show how we can use Henrich's model to positively shape the future. Until then!

# The WEIRDEST People in the World by Joseph Henrich

[MRM]

*Why are some countries rich and others poor? Does it depend on whether you, and your friends, and your grandparents (and their grandparents) married your cousins?*

That explanation sounds intuitively incomplete, but Joseph Henrich builds a mountain around it in [The Weirdest People in the World](#). He convincingly argues that Western European (plus North American, or shortened to WEIRD) people have a very different psychological makeup compared to other humans -- both other humans now, and other humans historically -- and those differences help explain the West's "success". His more provocative claim is the differences are most explained when, about 1,500 years ago, the Roman Catholic Church banned cousin marriage as part of a larger overhaul of family life.

I.

Did you know Westerns are WEIRD compared to others? Henrich frames his initial beef as one with the misapplication of academic psychology research. Unique human psychology findings from a Western university lab experiment don't necessarily apply to people born outside of the Western world, and certainly don't necessarily reflect innate human nature. Consider:

1. WEIRD (Western Educated Industrialized Rich and Democratic societies) are less likely, in lab experiments, to conform to a group (the Asch Conformity Experiment). Japanese subjects are about 1.6x more likely to conform (than WEIRD ones), Brazilian subjects about 1.7x more likely, and Zimbabweans about 3x more likely.
2. WEIRD subjects are relatively more patient, measured by willingness to wait for money (think an adult version of the "marshmallow test"). A Swede will forgo \$100 now to receive \$144 in a year. But globally, the average person requires \$189 in a year to give up \$100 now, and the far-end of the spectrum (Rwanda) requires \$212. (According to Henrich, the figures were adjusted for local currency and purchasing power.)

3. WEIRD subjects are less likely to lie in dice rolling experiments when lying is anonymous and also gets you more money. (The lying can be measured by an aggregate statistical analysis, i.e., whether higher numbered die were over-reported relative to baseline aggregate probability.)
4. In a hypothetical prompt, 80-100% of WEIRD subjects believe that their friend has no right to expect them to falsely testify in a trial so that the friend avoids legal punishment. That's about double the rate of India/China subjects, who are more likely to lie for their friend.
5. WEIRD subjects are 50%-100% more "analytical" rather than "holistic" in their thinking; for example, in a "triad test", WEIRD folks would pair a rabbit with a cat (two animals) instead of a rabbit with a carrot (rabbit eats carrot).

Compared to non-WEIRDers, WEIRD people are:

*Highly individualistic, self-obsessed, control-oriented, nonconformist, and analytical. We focus on ourselves, our attributes, accomplishments, and aspirations over our relationships and social roles*

Henrich's point is that *this isn't normal*; rather, we're the abnormal (weird) ones.

Henrich emphasizes these are differences, not value judgments, and of course only measured at the population level. Is it a *good thing* that WEIRD culture has people, on average, more trusting of strangers and less biased towards family members? Of course that depends on your definition of "good". And also on the current environment/context: In a strong state with a market-oriented economy and rule of law, there's big payoff (especially collective payoff) in those traits, even if it might be unwise to trust strangers if you find yourself in a different place or different time. It's understandable that our cultures didn't uniformly evolve "trust stranger" norms.

II.

So WEIRD people have very different cultures and personalities than others. But why? Henrich's answer is basically that we got lucky when the Catholic Church decided to ban cousin marriage and embark on a new, and different, family marriage program.

First, a quick example of his overall "cultural evolution" framework. In Henrich's model, different cultures develop different cultural norms, but within a dynamic process where the culture is also competing with other cultures, resulting in cultural selection

pressures (akin to survival of the fittest culture). Henrich recounts an illustrative New Guinea example:

1. New Guinea villages had a maximum of 300 people, about 80 of which were male. When villages grew larger than that, social ruptures tended to follow -- as though there was an “invisible ceiling” on the scale of human cooperation in New Guinea.
2. But one community, Ilahita, was an exception. Ilahita had a population of about 2,500.
3. Anthropologist Donal Tuzin explains the exception by reference to a unique cultural package within Ilahita, largely related to institutionalizing more cross-cleavages within that population. As described, a family might belong to “subgroup A-2”, which meant it interacted with a broader community of other A’s (not all of which were also “2’s”) **and** interacted with a broader community of other 2’s (not all of which were also “A’s”), meaning the entire community was interconnected, but in a relatively random way. This is intuitively familiar to us moderns, who have our family group, work group, social group, sports team group, etc., in each case with only some overlap. But maybe it was less common in historical human societies. (The anthropologist and Henrich discuss other cohesion mechanisms, like adolescent rituals, that I’m giving short shrift to.)
4. These cross-cleavages allowed Ilahita to scale to a higher population than other villages that were organized exclusively (or more-so) on family/clan grounds.
5. Why did Ilahita have these successful customs, such as cross-cleavages, but other villages didn’t? Apparently around 1870 the Ilahita leaders learned and copied some rituals from another successful group (the Abalam), but made errors doing so. Those copying errors turned out to work even better than the original customs, and thus proliferated. An accidental mutation, but one that happened to be useful.

### III.

So human history includes culture competition and cultural evolution. Still, as best as we can tell, most of human history was dominated by a “kinship” and clan-based approach to organizing society. Things like *who you are*, *what you do*, and *who you marry* all had a lot to do with your family background and the small community you were born into.

Here are some traits of the familial (or kinship) approach to culture:

1. Newly married couples set up residence near one person's (usually groom's) parents.
2. Inheritance of property after death is totally familial, usually via your father. (Admittedly inheritance is still mostly familial, but that's at least an option in modern society.)
3. "Corporate" (or collectivist) notions of responsibility for each family and clan. *"If someone kills someone from your clan (even accidentally) - their entire clan is responsible for paying blood money."* If they don't, *your entire clan* is responsible for a revenge killing.
4. Marriages: Often there are taboos *against* marrying a sibling. But then norms *in favor* of marrying a relative like a cousin (or second cousin). Henrich also flags the widespread popularity of "levirate marriage"; if a husband dies, then the wife *has* to marry one of his brothers or cousin-brothers, in order to maintain the familial link established by the original marriage.

None of the above is in any way universal, tribes everywhere differ, etc., but I take it as a rough picture of relatively common pre-modern practices. Some of these practices are still common today in non-Western places. Pre-Catholic takeover, European tribes mostly reflected the above: Kinship-based units collectively owned territories, provided social identities, adjudicated disputes, provided protection and health care, and arranged marriages. High-status European men married multiple wives.

How about now? 75% of global historical societies studied by anthropologists exhibit marriage to cousins or other relatives, yet WEIRD societies don't do that. 72% trace lineage through only one parent, while WEIRD societies trace lineage through both parents. WEIRD societies are part of the only 8% of all studied societies with nuclear families, and part of only 5% where newly married couples set up a separate household. Circa year 2020, you can move to NYC for a job and then mate up, and then do whatever you want next (move to the suburbs, move closer to your parents, move closer to her parents, move to California). It's easy to appreciate how this family freedom is probably historically unique, but is it historically meaningful?

IV.

According to Henrich, it was indeed Christian religious peculiarities, which ultimately impacted family formation, that explain how we got from the traditional kinship culture to our modern Western culture. The first key development was universalizing religions,

which include Christianity, Hinduism, Buddhism, and Islam. These universalizing religions had, in certain settings and certain respects, advantages over the previously more common local god approach. They started the path of universal rules or principles that apply to everyone equally. They also tend to believe in contingent afterlives and free will, other factors that prepared the Western world for its cultural evolution.

Really, though, it was the Roman Catholic's family and marriage program that changed the game.

Henrich's chart of the Catholic marriage program starts in year 305 when Synod of Elvira forbid men from marrying either (i) a sister of their dead wife or (ii) their daughters-in-law. I don't know who Synod of Elvira is, or whether he was particularly important, or if such sister-of-dead-wife / daughter-in-law marriages were so frequent so as to be historically significant. The point is the trend. By year 500-600, first cousin marriage was officially outlawed by Roman Catholic popes and authorities; the Eastern Church, which lagged in both time and intensity of kinship reform, made that official in year 692. The marriage restrictions gradually expanded. By about year 1050, the ban reaches sixth cousins or "*marriage with kin as far back as memory goes*". (Aside: In 1983, John Paul II loosened incest restrictions and allowed second cousins and more distant relatives to marry.)

The family and cousin marriage prohibitions forced folks to search further and wider for potential mates:

*By roughly 1000 CE, manorial censuses confirm that peasant farming families [in France] lived in small, monogamous nuclear households and had two to four children. Young couples often formed independent neo-local households, sometimes moving to new manors.*

By the late Middle ages, Europeans had later marriage ages (mid-20s), a relatively large percent of unmarried women, smaller families and lower fertility, and frequently worked during a premarital labor period. Henrich contrasts that cultural European

landscape to the cultural patterns seen, for example, in China around that time, which still reflected the traditional kinship cultural practices.

V.

Large societal (and, later, psychological) consequences followed the disruption (really deterioration) of Europe's family culture. Simply put, people intermingled more, and family bonds weakened. Henrich amasses chart after chart to show that we can see current data of strong correlations between personality traits and background family structures. Countries with *more* cousin marriage and more familial kinship cultures (i.e., non-WEIRD places) are *less* individualistic, *more* conformist, *more* shame-ridden (but fewer feelings of guilty), *less* trusting of strangers, contribute *less* in (stranger) group projects, voluntarily donate *less* blood, etc. (this list could go on for pages).

Henrich also assembles statistical evidence showing a dose-response effect: The longer a population or region was exposed to the Western Church's marriage family program, "*the weaker its families and WEIRDer psychological patterns are today.*"

Of course the number of correlations don't prove Henrich's theory of causation; it's possible that European societies had a different secret sauce that propelled them along their path, and *that* secret sauce (directly or indirectly) causes our modern psychology differences. But his theory is facially plausible. Europeans may have stumbled across a religious mechanism that started weakening familial bonds, gradually at first, but slowly increasing (and resulting in) non-familial bonds, relationships, and institutions like guilds, charter cities, monasteries, universities, and quasi-markets. None of these replacement institutions are particularly impressive or important *at first*. But they, alongside the unrelenting advance of the marriage program, further contribute to additional weakening of kinship-based institutions. Eventually the quasi-market institutions develop into even more institutionalized markets. The culture and norms encouraged by the new institutions are different from those encouraged by kinship institutions, such as universalizing treatment, positive sum worldviews, fairness towards strangers, and reciprocity. Collectively, they build a sort of interpersonal prosociality. The interpersonal prosociality norms made larger cooperation, and larger

urban locations, more viable. These effects continue building on each other; weaker family bonds leading to stronger non-family institutions, leading to larger cities, which in turn further solidify a market approach and weaken family bonds. By about 1200, Belgium, the Netherlands, and Italy started urbanizing; England urbanized later (starting in year 1500), with an overall European urbanization quadrupling (3 percent to 13 percent) from year 800 to 1600. If the industrial revolution represents a fundamental change in the human condition, for Henrich, humanity needed something (anything?) to jump us off the familial relationship track and get us onto the prosocial one. The Catholic church and its family marriage program just happened to be the ticket.

VI.

Henrich points to the WEIRD “clock-time mindset” as one example of the importance of cultural evolution on our psychological differences. WEIRD folks are obsessed with concepts like hourly work efficiency, being “on time”, and not “wasting time.” Henrich argues this is a relatively new psychological development that did not exist before the High Middle Ages. Another fascinating example is the Western legal tradition of investigating “*universal legal principles, categories, or axioms from which all specific [laws] could be derived.*” According to Henrich, this search for universal legal principles came from an application of WEIRD personality when re-reading Roman case law. It was another mis-copying of sorts, since universal principles did not exist in the Roman legal sources, but Europeans, indoctrinated by hundreds of years of prosociality and universalizing, assumed they must exist.

In Henrich’s telling, Martin Luther’s protestant reformation is perhaps the final booster shot that Europe needed, explaining why Europe (and not others) got modernity, the Enlightenment, the Industrial Revolution, representative government. By this point of the story, things are very cyclical: Europe was simultaneously predisposed to accept *something like Protestantism* when it came along, and *Protestantism further* pushed Europe along its developmental path. Relative to Catholicism, Protestantism further individualized our culture and psychology by emphasizing things like reading, self-education, intentions, hard work, and self-discipline. Literacy rates subsequently grew



faster in Protestant European countries (such as Britain, Sweden, and the Netherlands) than Catholic European countries (Spain and Italy). If Europe was WEIRD, Protestant Europe became WEIRDer: More individualistic, more rational, less familial, and, later, more Educated, Richer, and Democratic (wEiRD)

*By instilling thrift, patience, and an internalized work ethic while at the same time requiring literacy and encouraging schooling, Protestantism had psychologically prepared the rural [European] populace to participate in and fuel the Industrial Revolution. Evidence from the 19th century...shows that compared to Catholicism, early Protestantism fostered higher literacy rates, greater incomes, and more engagement in manufacturing and service industries. Politically, Protestantism probably encouraged the formation of democratic and representative governments.*

## VII.

According to Henrich, Jared Diamond's *Guns, Germs and Steel* explain a lot of dispersion amongst societies and global-level inequality up to about year 1000. But afterwards...

*If a team of alien anthropologists had surveyed humanity from orbit in 1000 CE, or even 1200 CE, they would never have guessed that European populations would dominate the globe during the second half of the millennium. Instead, they probably would have bet on China or the Islamic world. What these aliens would have missed...was the quiet formation of a new psychology.*

Does Catholicism, family marriage policy, protestantism, and the resulting cultural psychology changes explain Europe's post-1200 trajectory? Does our propensity to marry our cousin (and that of our great-great-great-etc. grandparents) represent a fundamental break in our history? Henrich's account seems facially plausible, even if it needs further testing and scrutiny from sociologists, anthropologists, statisticians, and historians. And even if this particular theory is shown invalid, it may still endure by drawing attention to "cultural evolution" and the interplay between cultural institutions

and underlying human psychology. Did the European cultural “advances” make Westerners lonelier and sadder while simultaneously richer? Did the human experiment ultimately benefit from the change, or will {nuclear annihilation, global climate change, or AI singularity} ultimately show the folly of breaking down kinship bonds? Henrich doesn’t go there. But such is our luck, good and bad, to be born into the WEIRD lineage. We may eventually realize our psychological approach to the world, just like our physical bodies themselves, resulted from the combination of luck, mutation, and incremental change that underpins most evolved outcomes.

\*\*\*

#### Endnotes:

1. I should add a global cautionary note that Henrich is often more measured in his conclusions than my summary indicates. For the most part, this isn’t because I disagree with Henrich, but rather want this idea to be easily digested without bogging this summary down with endless qualifications. But also to say that if your primary objection to something is “this seems to be overstating the case”, that’s probably an objection against my summary more than Henrich’s work.
2. Henrich devotes some pages to discussing polygamy, arguing that it was probably the natural state of human sexuality and monogamy is deviant. There is already a blog post that Henrich “Gets polygyny wrong.” Although I’m by no means an expert, my understanding of the issue (the prevalence of polygyny v. monogamy in early civilization and pre-civilization) is also softer than Henrich’s, primarily influenced by Sapolsky’s *Behave* and Geoffrey Miller’s *The Mating Mind*.  
<https://traditionsofconflict.com/blog/2020/9/8/the-weirdest-people-in-the-world-gets-polygyny-wrong>
3. Why did the Catholic Church care about this kinship marriage stuff, and other religions didn’t? Henrich has some pages on this, although he also seems to treat it as a stochastic event (remember the bit earlier about cultural mutations for the New Guinea tribes). I don’t think explaining the origin is actually all that important or central to his argument.

# The World Imagined by Hendrik Spruyt (2020)

If you read my book review of *The Sovereign State and Its Competitors*, you can guess what this book is about. A sovereign, territorial state is an institution which claims a monopoly on violence and justice within certain lines on a map and no authority elsewhere. States are often associated with a particular cultural community, the nation. The international system of sovereign, territorial states has become so dominant that it's hard to conceive of anything else. This has not always been the case.

In *The Sovereign State*, Spruyt looked at how the sovereign state developed in Western Europe from the interaction of feudalism and urbanization, and how other institutions arose in the same circumstances but were ultimately outcompeted. In *The World Imagined*, Spruyt shifts his focus later in time and broader in geography. He looks at three Early Modern international societies: the Chinese tributary system, the Ottoman / Safavid / Mughal Islamic Empires, and the galactic polities of Southeast Asia. Each of these international systems interacted with colonial Europe, with its already established sovereign states. Although this interaction ultimately led to the dominance of the sovereign state system, each developed in response to the other.

This book isn't as good as *The Sovereign State*. Spruyt is farther from his area of expertise, so his knowledge is less deep. Instead of having good, specific examples for each of his claims, he is much more likely to quote another scholar. He also spends more time debating other schools of political philosophy. Although I don't really care about why the early English school is wrong (for example), I wouldn't really mind it - except that it leaches into his history. There are multiple places where I think that the stories he tells are designed more to support his position in academic debates than to accurately represent the past. Despite these criticisms, it has led me to some interesting thoughts. My responses to Spruyt's ideas are in {}.

---

Like *The Sovereign State*, this book begins with a literature review that places Spruyt's ideas in the broader academic context. If you don't know or care what ontology is, you can safely skip this section.

[Ontology is the study of what it means for something to exist. Do the pen I am currently holding, the number twelve, and the country of Poland all exist in the same sense?

Should some notions of existence be prioritized over others?]

Current political theory is dominated by disagreements between Realists, who think that states primarily pursue their own material ends, and Neoliberals, who think that states primarily try to cooperate with other states to promote their values (from democracy to free movement of capital to the immutability of borders). Both theories assume that international affairs is the study of the interactions between states. Since there is no world government, there is anarchy between the states.

Spruyt is especially critical of the positivist / empiricist trend in the social sciences. Partially I think that this is him defending his turf against encroachment from the harder sciences. But he does make several compelling points. In their search for general laws, empiricists tend to try to understand all historical systems using modern ideas, instead of trying to understand the system as the people who lived it would. Empiricists also tend to prioritize measurable explanations such as relative economic or military strength over unmeasurable explanations such as ideas of honor and shared ritual identities. The empiricist approach is especially likely to go astray because political systems are not things outside of ourselves that we can impartially measure. Instead, the political system is created by how we collectively imagine the world:

"Our understanding of the social world is co-constitutive of that world. That is, our theories, our beliefs of what the world is or should be, simultaneously create the social world which is the object of study." [43]

I will include two more quotes from this section.

The first describes how Spruyt understands the Westphalian system of sovereign, territorial states:

"The Westphalian system enshrines the principle that the governing authority should be defined territorially, with borders demarcating the sphere of domestic politics. Within this sphere, individuals ideally identify themselves as a nation, the imagined community whose identity is linked to the territory in question. Territorial state and national identity are thus fused, to define the in-group in which individuals find purpose and mutual obligation. Beyond the borders, interactions between states occur in the realm of anarchy. Nations in other states thus constitute the natural and oppositional "Others" or outgroups. In the Westphalian system inclusion and exclusion are thus spatially defined.

Historical reflection suggests that this perceived ontology of the contemporary state system is not a necessary and natural reflection of a material phenomenon. By assuming that all international relations can be understood as a set of interacting

elements under anarchy, social creations have been interpreted as if they were brute fact." [31-32]

The second describes Spruyt's definition of an international society, without assuming states:

"I define international societies as a set of polities that share foundational collective beliefs - a collective imagination regarding the nature and purposes of political and social organization. These beliefs thus serve to define the nature of the political order and legitimate authority. Second, collective beliefs demarcate the boundaries of the political community. Who is rightfully considered a member? Who is deemed an outsider? Third, they contain formal and informal principles that govern the interaction between members in the system as well as with polities outside the system. Fourth, collective beliefs influence the status of polities that are not part of that international society." [52]

We now turn to our first alternative international system.

## Chinese Tributary System

China was ruled by an emperor. He claimed the authority to rule "All under Heaven". China was the dominant power in its neighborhood. Whoever ruled the North China Plain had access to far greater resources than any of their neighbors. Despite this economic and military dominance, there were limits to their power. China could not militarily control the ends of the earth.

Nevertheless, the Chinese tributary system was remarkably successful at reducing interstate warfare. There were only two interstate wars in East Asia in the 500 years preceding the Opium Wars between China and Britain.

{ I am much less impressed by this claim than Spruyt is, because it only looks at wars between states. War within states also existed - especially in Japan. But more importantly, most of China's neighbors weren't states. Spruyt notes that the nomadic peoples who lived to the north and west typically rejected the Sino-centric system. This is very significant: two of the last three dynasties to rule China were founded by invading nomads. Spruyt doesn't even seem to be aware of the hill peoples to the southwest, who were explicitly anti-state. Uncommon wars between states is much less impressive when states are uncommon. }

Instead of military power, China insisted on its cultural superiority. Nearby states (especially Vietnam, Korea, & Japan) were expected to recognize this superiority and adopt Chinese practices, especially their calendar, writing system, and Confucian civic rites. Other states would send diplomats to the emperor and ritually submit to him. The emperor could claim that his rule extended over China's neighbors as well. The diplomats would give tribute to the emperor and he would bestow gifts to them in return. The value of the gifts from the emperor often exceeded the tribute given. States that ritually submitted to the emperor would also be given trade privileges in China and a guarantee of no military interventions. They could use the legitimacy granted by the emperor in their struggles against internal rivals.

The ritual submission underlies much of Chinese culture, not just interstate politics. People understood their social status through prescribed ceremonies between unequal status: between heaven and earth (represented by the emperor), between the emperor and governors, between state officials and commoners, and between ancestors and children. The inferior would give honor and tribute and the superior would bestow gifts.

Chinese monumental architecture is designed around concentric circles (representing Heaven) and squares (representing Earth). The emperor would ensure the continuation of the seasons by harmonizing Heaven and Earth at these sites. Entire cities were laid out as models of the cosmos.

Chinese maps similarly reflect the cosmological-moral world, with China at the center. When European maps reflecting the physical world became available, China did not adopt them, because maps fulfilled a different role in society.

{ Somehow I think that someone would have a harder time navigating the cosmological-moral world using a Chinese map than the physical world using a European map. }

The emperor would also write histories and regulate knowledge to further his claims of superiority. Chinese histories regularly describe how foreign rulers are embracing Chinese culture (although histories in other languages would not). The emperor claimed to be able to predict signs in the heavens, so unauthorized dabbling in astronomy was outlawed. Ritually performed and recorded obedience was more important than actual obedience. The emperor's power lay in his ability to get other people to treat him as emperor.

The tributary system was a framework for understanding interaction, rather than a rigid system of rules. Different people within the system could use this framework to promote their own interests. In Korea, potential heirs of a king would persuade the Chinese emperor to recognize their claim, which helped to establish their legitimacy within

Korea. In Japan, the emperor declared a ritual superiority over local rulers (and, in theory, over foreign powers), despite having little material power. Ming histories record the "benevolent gifts" offered to Mongol hordes who had come to "pay tribute" to the emperor, even when the material balance of power did not reflect the ritual one. The Manchu rulers of the Qing dynasty fulfilled the appropriate Confucian rites to establish their legitimacy, while still promoting their own ethnicity within the government. As long as the actors fulfilled the appropriate roles in the system, they could use the system as they wished.

{ I expected more from this section. East Asia clearly had multiple worldviews: the Confucian agrarian state, the anti-state societies of the hills of southern China and southeast Asia, the extremely flexible sovereignty of the nomadic confederations on the steppe, and the Buddhist theocracy in Tibet. These worldviews competed and overlapped for hundreds, if not thousands of years before the Western powers arrived. Spruyt mentions the idea that the Qing dynasty had learned to successfully combined or operate within multiple worldviews, which allowed them to expand into areas traditionally not part of China: Manchuria (the Qing homeland), Inner Mongolia, Tibet, Xinjiang, and Yunnan, but he hasn't developed the other worldviews enough to show how they were synthesized. }

Spruyt then describes and criticizes the traditional narrative of the interaction between the tributary system and the sovereign states. The traditional narrative emphasizes the incompatibility between the two systems of international relations. The Chinese proved to be unable to adapt ideologically and institutionally and so declined relative to the West. Japan, which was not as closely tied to the tributary system as other East Asian states, was able to adapt and redefine itself in the terms of the Western powers.

Spruyt's response is that China did adapt to Western ideas and modernize its institutions after the Opium Wars. This process was resisted by conservative factions within China, and hindered by unequal treaties with Western powers and later Japan. The reason why we think that China didn't adapt is because the West wanted China to remain an Other with an illegitimate government, so it could be subject to imperialism.

{ I think Spruyt missed a main point of the traditional narrative. China's ideological and institutional stagnation was less important than its economic and military stagnation. The real or imagined offenses to the British ambassador when both sides refused to ritually submit to the other are less important than the fact that Britain's military was stronger than China's, despite having a tenth the population. Similarly, Spruyt's description of Japan's transformation focuses on how they mastered Western diplomacy and were able to convince China and European powers to sign treaties as equals. He doesn't mention that in 1853, Japan couldn't deal with 4 American ships, while in 1905,

Japan crushed the entire Russian navy.

This is one of the times that Spruyt's focus on the academic debate undermines the history. Spruyt thinks that we should understand history in terms of collective imagination, not realist power calculations, so Spruyt ignores the power balance. This refusal to think in terms of power makes his narrative incomplete. Why didn't China go on (a lot of) voyages of discovery? Why didn't China industrialize? Why didn't China institute an effective military draft? These questions are not answered by Spruyt, so he does not provide an alternative to the traditional narrative. }

Some modern Chinese scholars have suggested that we look at the Chinese tributary system as a more peaceful and harmonious alternative to the Western state system. Scholars among China's neighbors are skeptical because China has spent a lot of time conquering its neighbors, not just harmoniously assimilating them.

{ Spruyt's discussion here could be much stronger if he considered whether there are any relevant modern examples. I can think of two:

The most obvious one is the "one country, two systems" in Hong Kong. Hong Kong (and the British) agreed that Hong Kong would be officially under China's control and the People's Republic of China agreed to respect Hong Kong's autonomy and legal system. China has actually been gradually forcing its institutions onto Hong Kong, against Hong Kong's will.

Tibet and Xinjiang were colonial projects of the last Chinese dynasty. Currently, both of them are officially autonomous regions within the People's Republic of China. However, the Tibetans and Uighur who live there are significantly less free than people elsewhere in China.

In both cases, we have something sort of like the tributary system: smaller polities near China officially recognize Chinese suzerainty, while China officially recognizes their autonomy and local traditions. Except, unlike the historical tributary system, the People's Republic of China actually imposes its institutions on the smaller polities using military force. These modern examples show that we should be skeptical when the People's Republic of China promises to respect the autonomy of territories officially under their rule. }



## Islamic Cultural Community

During the Early Modern Era, the Islamic world was dominated by three Turkic Empires: the Ottomans in the Mediterranean, the Safavids in Persia, and the Mughals in India.

The Turks were originally nomads from the steppe. They first entered the Middle East as mercenaries for the Islamic caliphate, which itself had been built by nomads from the deserts of Arabia. During the conquests of Genghis Khan, many Turkic tribes allied with the Mongols -- and those that didn't were annihilated. Unsurprisingly, steppe nomads were not organized as states:

"The political organization of the steppe people was confederate rather than hierarchical in any strict sense. Individuals identified with the tribe as the critical marker of differentiation from others. Periodically, when waging war, tribes would create larger entities on a loose confederate basis with war leaders, khans, giving direction to the affiliated tribes. The success of the war leader conferred legitimacy. The lack of success demonstrated the converse, loosening the obligation to follow the leader." [207]

The Mongol Empire did not last for long. The steppe practice of dividing your inheritance among your sons works better for flocks of sheep than for transcontinental empires. Timur (Tamerlane) was the last to try to reconquer all of Genghis Khan's empire. The Turkic Empires all claimed to be heirs of the world conquerors: Alexander the Great, Genghis Khan, and Timur.

{ It's interesting the three people who appear on this list. }

Once the Turkic Empires conquered lands filled with farmers, they adopted the administrative practices, and titles, of the previous rulers. They were not only khans (a Mongol term) and sultans (an Arabic term), they were also the Kaysar (Caesar) / Basileus of Byzantium or the Shah of Persia. Nomadic practices that made less sense in a settled society were adapted. The sultan led the annual military campaign, instead of being present during the continual raids. To avoid the problems with dividing inheritance, when the Ottoman sultan died, one of his sons would kill all of his brothers.

Islam was extremely important to these three empires. Each also claimed the title Caliph, ruler of the Muslim community.

{ Spruyt downplays some of the evidence he presents here because he is worried that people will use it to support the "Clash of Civilizations" narrative that sees the Islamic world and the West as fundamentally incompatible. }

Unlike Christianity (but like Judaism), Islamic scripture contains a legal code: Shari'a. Law is not something that is chosen by our representatives in a legislature or constitutional convention; law is something that is given by the divine. Along with the law contained in the scriptures, there are also centuries of commentaries on the law, much like how American constitutional law contains both the Constitution and the Supreme Court cases about the Constitution. There are disagreements in the commentaries: Sunni Islam alone recognizes four different schools. But it is clear that political leaders should be enforcers of the divinely mandated law, as interpreted by religious scholars, not creators of the law.

The Qur'an divides the world into Dar al-Islam and Dar al-Harb, the House of Islam and the House of War. Some scholars argue that there should also be a third House, which is at peace with Islam, consisting either of tributary states to Muslim leaders or of states that have signed treaties with Muslim leaders who allow Muslims to live peacefully. But this third House was not always recognized, especially in the early modern era. Muslim states often considered themselves continually at war with all non-Muslim polities. Cease-fire agreements that lasted for less than ten years were allowed, but not peace treaties.

{ I think it is more accurate to interpret Dar al-Islam as the territories ruled by Muslim law, not the territories where Muslims are a majority. Spruyt deemphasizes the division of the world into two parts by showing how Muslim rulers ruled over significant numbers of non-Muslims. That does establish that the dichotomy doesn't advocate for genocide (like the people who want to create a caliphate today do), but it doesn't mean that they didn't divide the world this way. }

{ How long was the nuclear agreement that wasn't a treaty between Iran and Obama (and others) supposed to last? Ten years. }

This ongoing religious war to expand Islam (jihad) was merged with the older Turkic practice of continually raiding your neighbors for horses and slaves (ghazi). Throughout the Early Modern Period, the Turkic Empires were continually at war with their non-Muslim neighbors.

Within the Islamic Empires, there was significant religious and cultural diversity:

"Karen Barkey describes the Ottoman policy to deal with this diversity as "separate, unequal, and protected". Religious boundaries were recognized and maintained. Ottoman rulers decreed dress codes, marriage, and other markers for distinguishing the non-Muslim groups from Muslims, although the degree of enforcement is open to

question. The underlying idea revolved around maintaining religious and ethnic boundaries and giving them protected, dhimmi, status. Groups would be self-policing with their leaders acting as conduits between the Ottoman state and the minority group membership.

Tolerance and protection, however, did not mean equality. In law courts, evidence from Muslims was privileged over that of unbelievers. Tax rates for non-Muslims were also higher. Muslims paid the lowest rates, dhimmis (Christians and Jews) resident in Muslim polities paid the jizya at an intermediate rate. Foreigners, those from Dar al-Harb, paid the highest rate as harbis. Demonstrative public behavior such as giving way to Muslim pedestrians, and more importantly prohibitions against worshiping in proximity of mosques, all served to reinforce the secondary status of Christians and other non-Muslims.

The effect was to create material incentives to convert to Islam. At the same time it also induced local elites to maintain these boundaries, since they could exercise patronage and control over the group members, as conduits with the center. The later millet system, which formally recognized the extrajudicial character of non-Muslims, only formalized much earlier practices." [198-199]

The Mughals were more tolerant of religious diversity than the Ottomans. A majority of their subjects were polytheists. The Safavids were less tolerant of religious diversity. Especially in their later years, they pursued policies that caused the conversion of the majority of the population from Sunni Islam to Shi'a Islam, the religion of the rulers.

{ The Turkic Empires were not tolerant by modern standards. One of the reasons I included the lengthy quote is to point out that this system could be called apartheid. Apartheid is better than genocide, which was not uncommon in Early Modern Europe, but it hardly is tolerance.

This is even more clearly seen from the devshirme system. The Ottomans would take boys from Christian households as slaves and force them to convert to Islam. Later in their life, they could rise to high ranking positions in the military and administration of the empire. Spruyt repeatedly uses this as an example of inclusivity. I definitely would not. }

European observers of the Ottoman and Mughal courts were impressed by how tolerant they were, which helped to influence the development of liberalism in the West.

{ This is the second time recently that I've seen a claim that liberalism has non-European roots, the other being Mann's claim in 1491 that freer Native American societies influenced American liberals who influenced Europe. This is a bold claim and

should be backed up by serious scholarship. The idea that societies should be free and democratic is easily one of the best ideas to spread out of Europe, so everyone wants to claim it as their own. The debate over whether Christianity was necessary for the birth of liberalism or hostile to it has gone on for centuries. The debate over non-European influences on liberalism deserves no less scrutiny. Spruyt's evidence is the sentence:

"Seventeenth-century scholars such as Locke and Voltaire and European ambassadors remarked on the tolerance of these dynasties compared to European Christian rulers. For all the contributions of the Enlightenment, tolerance was neither a European invention nor a European monopoly." [206]

There is no reference to a more thorough investigation or even to the specific remarks. This is sloppy scholarship and makes me wonder if it is included, not because the author has good evidence to support this claim, but because he likes the political conclusions it implies. }

The three Turkic / Islamic Empires had shared cultural and religious understanding that they constituted a single international system.

The Safavids and Mughals had many reasons to go to war: they both claimed to be world conquerors, the Safavids ruled the Mughal homeland, they both claimed to be the lord of all Muslim, and they belonged to different sects of Islam. But because of their shared cultural heritage, they only went to war occasionally and usually recognized each other as equals.

The Safavids and Ottomans were less amicable. The Ottomans considered the Shi'a Safavids to be infidels, so they were part of the House of War. There was continual war between these two empires for almost the entirety of the Safavids' reign. Despite this war, there were extensive trade and cultural connections. Persians were always allowed to go on the hajj to Ottoman controlled Mecca. Even though they were enemies, they had similar understanding of the foundations for politics and society.

Both the Safavid and Mughal Empires collapsed to internal challenges before the sovereign state system was imposed by the West. The Ottomans adapted to survive. The Ottoman policy of continual war with their non-Muslim neighbors lasted as long as they were still expanding into Europe. After the failed Siege of Vienna in 1683, they started losing ground to European powers and began to adapt to the European state system.

{ The Siege of Vienna was broken by the Winged Hussars of Poland. Yes, they did wear wings into battle. The largest and most consequential charge by fully armored cavalry

occurred almost 200 years after the end of the Middle Ages. }

Delineation of borders was not a problem: first with the Safavids in 1639 (Treaty of Zohab) and then in Europe in 1699 (Peace of Carlowitz), although frontier zones with lower levels of conflict existed in both places. Ottoman maps were geographic, like European maps, but with Istanbul at the center.

Modern diplomacy, with "permanent missions, extraterritoriality, reciprocity, and the gathering of intelligence" emerged among the Italian city-states in the fifteenth century [226]. From the Muslim perspective, this looked more like exchanging hostages than diplomatic immunity. The Ottomans did not participate until the 1830s. After this, they were part of the European state system. The Ottoman Empire was not invited to the Congress of Vienna in 1815 at the end of the Napoleonic Wars because the Holy Alliance (Russia, Austria, and Prussia) found them barbaric. They did participate in the Treaty of Paris in 1856 ending the Crimean War.

Before then, there were numerous temporary and semi-permanent emissaries between the governments. An agreement with France was signed as early as the 1500s (both really hated the Habsburgs), but most European countries didn't diplomatically engage the Ottomans until the 1700s. Even then, both sides viewed the other with continual mistrust. European powers repeatedly took territory from the Ottomans until the nation-state of Turkey emerged after WWI.

{ Spruyt shows that Europeans thought of the Ottomans as Others and that the Ottoman Empire lost territory to European states, but I do not think he shows causation. While the boundaries of Western European states haven't changed much since 1700 (with a few notable exceptions), the boundaries of Central and Eastern European states certainly have. European powers that lost major wars were divided up among their neighbors, including the Swedish Empire, Poland (multiple times), and the Austro-Hungarian Empire. Even if the Ottomans had been accepted as Europeans, they probably still would have lost their empire. There is a case to be made that the Arab protectorates were formed because Arabs were Others, but Spruyt doesn't make it. }

The bigger ideological problem for the Ottoman Empire, and all multi-ethnic empires, was nationalism. This can most clearly be seen in the Greek revolt of 1830, which established a small independent state of Greece, while the majority of Greeks still lived under Ottoman rule.

{ Nationalism is the idea that people can be divided into linguistic / religious / cultural groups. Each nation has the right to a sovereign, territorial state which includes everyone in their nationality.

While there are some hints of this idea as early as the Hundred Years War between France & England in the 13-1400s, nationalism wasn't fully articulated and widespread until the 1800s. In Western Europe, the logic worked the other direction: states created nations instead of nations creating states. The uniform culture of France, for example, was created through hundreds of years of public education, national historiography, religious genocide, and military drafts. Peasants had to be made into Frenchmen [Weber, 1976].

Nationalism is an extremely problematic idea, especially in a multi-ethnic society like the Ottoman Empire. It led to large forced migrations between Greece and Turkey after WWI. It caused the Balkans to become, well, balkanized. It caused the newly nationalistic Turkey and the new Arab states to repeatedly genocide the Armenians and Kurds.

This isn't something unique to the Ottoman Empire: Europe only has nation-states because of centuries of genocide and forced migrations. While there were some attempts to make political borders follow cultural borders, more often, people in the wrong state were relocated, assimilated, or killed. Nationalism is still a major cause of forced assimilation and rebellion. }

The Ottoman Empire did not react passively to the encroachment of the West. It adapted and reformed its institutions until it became the sovereign, territorial nation-state of Turkey.

{ This is why the Ottoman Empire was the first non-European / non-Christian country to be included in the sovereign state system and why Turkey survived colonialism better than almost all other non-European countries. }

## Galactic Empires of Southeast Asia

Spruyt's main point of these chapters is that, despite their diversity, Southeast Asian states had a shared political understanding of the world and so could be considered a single international system. This shared understanding developed without the unifying influence of a dominant power or a shared religion.

{ I do not know enough about the historical states of Southeast Asia to evaluate this

claim.

I have read one other book on Southeast Asia recently, Scott's *The Art of Not Being Governed* about the anti-state peoples of the hills. Spruyt is completely unaware of them. He even quotes another scholar who refers to the frontiers of states as "broad zones of emptiness" [257]. The frontiers were empty of states and written records, but they were certainly not empty of people. The single international system of Southeast Asia that Spruyt describes only applies to the people living in the lowland states. }

Both Hinduism and Buddhism heavily influenced Southeast Asia, and Islam spread into Malaysia and Indonesia starting in the 1400s (only a few decades before the European traders and colonizers). Despite having different religions, the states had a shared vision of cosmological order that they reproduced politically.

To understand Southeast Asian states, we need at least a basic understanding of Indian religious cosmology and sacred numerology. This is summarized in a mandala diagram. I will describe a characteristic version. Other variations, both in the diagram and its political realization, exist.

"Mandala" itself is a combination of core (manda) and enclosing (la) parts. A mandala is a set of concentric circles, with 8 lines radiating out from the central circle in the cardinal and ordinal (NW, SE, . . . ) directions. These circles are surrounded by several concentric squares. The mandala is also supposed to be three-dimensional, so two-dimensional depictions of them typically have additional details to convey height.

The central circle, from which the 8 lines radiate, represents Mount Meru, the home of the gods. It is surrounded by seven rings of ocean and seven ranges of mountains. You must pass through these mountains and oceans on the way to enlightenment. Beyond the outermost mountain range is an ocean with four continents, in the cardinal directions. People live on the southern continent.

The mandala scheme was repeated in many aspects of Southeast Asian states.

The capital city was designed as a mandala. At the center of the capital, a tall temple or palace, hopefully built on a hill, stood as a representation of Mount Meru. The king stands at the center, on the axis between earth and heaven, creating order out of chaos, and causing the world to turn.

The geography of the state was also configured using this pattern. Along with a capital, there should be four core provinces and four outer provinces. Local rulers valued the numerology over their material interests, even when the central authority was weak or

non-existent. The Maluku island chain, now in central Indonesia, always had four kingdoms. The two most powerful were always rivals, on opposite sides of cosmic dualism. Nevertheless, they were happy to cooperate in maritime trade. One side allied with the Dutch and the other with the Portuguese. After the Dutch subdued Jailolo, one of the four kingdoms, they tried to incorporate it administratively with other islands. This led to a revolt, even on the islands which had been rivals with Jailolo. Maluku had to consist of four kingdoms.

The ruler of each province would create his own capital using the mandala pattern, leading to a nested political structure. "Galactic" kingdoms refers to how lesser political replicas revolved around the central core. Power radiated out from the king, diminishing in radiance the farther you are from the center.

{ This is a model of a solar system, not a galaxy. Galaxies do have a central object (a supermassive black hole), but it is not radiant and contains only a small fraction of the total mass of the galaxy. }

The king had an inner circle of four inner advisors and an outer circle of four advisors. Each direction corresponded to a particular advisor. For example, the military advisor had authority over the south province.

The mandala system was understood more as personal ties between the king and his subordinates, not as the kingdom's ties with the land.

{ Scott has a better explanation for why Southeast Asian states emphasized ruling people over ruling land. States want to produce as much as possible, and before the Industrial Revolution, this means as much agriculture as possible. When there are more people than are needed to farm the land, then states focus on ruling land. When there is more land than the people can farm, then states focus on ruling people. The goal of war is not to seize territory, but to seize captives to bring back to the capital. Early Modern Southeast Asia had a low population density, so it was focused on ruling people. The first states in Mesopotamia similarly dealt with low population densities, and focused on ruling people and seizing captives, rather than conquering land.

This got me wondering when the transition from ruling people to ruling land occurred in the West. And I realized, to my surprise, that I already knew, because the story is still being told.

When the Babylonians conquered Jerusalem, they took almost everyone captive to Babylon. A generation or two later, the Persians conquered Babylon. Darius, king of Persia, decreed that the Jews could return to their homeland and rebuild their temple



and the walls of Jerusalem. The Persians (and almost everyone after) preferred to rule their subjects in their own homelands, according to their own traditions.

Spruyt would not like Scott's explanation because it shows how ideology follows from material conditions. Spruyt prefers to show how material conditions follow from ideology. }

A king was expected to rule according to dharma. While this could be translated as 'virtue', the meaning is different from the virtue expected from a divinely appointed righteous European king. Because the gods reward dharma, success implies that a person has dharma, perhaps accumulated in past lives. Control of ritually significant sites and performance of the pomp expected from a king showed accumulated dharma and inherently make the king legitimate. Correct descent is not as significant. Many Southeast Asian languages don't even have a word for 'usurper', not because usurping rule was uncommon, but because the new ruler automatically became legitimate.

Kings did not recognize each other as equals. There could be only one central capital around which the rest of the world revolves. Major kings would try to assert their dominance over all other surrounding kings, while lesser rulers would recreate the pomp of kingship as much as possible and look for opportunities to assert their independence or dominance.

Natural borders, like mountains or rivers, were sometimes recognized, but artificial borders were not. The first people in Thailand to encounter European maps thought that there were vertical planes dividing the countries. Frontiers were more often regions where the diminishing power radiating out of multiple great kings overlapped.

The combination of great kings refusing to recognize each other, many local power centers that symbolically mimicked the capital, and the fact the success automatically granted legitimacy made Southeast Asia extremely politically unstable. Wars between major states were common and could be genocidal. The focus was usually on taking captives and forcing assimilation more than on killing your enemies, although some island societies maintained headhunting traditions. The succession of kings was contested more often than not. There was a continual struggle between centralization efforts by the king and the periphery's desire for autonomy.

Islam began spreading into Southeast Asia in the 1400s, followed a few decades later by the Europeans. In most places where Islam is predominant today, Islam arrived first by the sword, then through the law, then with mass conversions. This pattern did not hold in Southeast Asia, where Islam spread through maritime trade networks. The mystic Sufi branch of Islam spread first. Islamic law did not become significant until the

1800s. Islam thus initially added to, but did not replace, older political frameworks. Islam also only spread readily in maritime Southeast Asia (Indonesia & Malaysia today). The mainland states continued to follow Indian religions.

The European colonial powers also initially adapted to local norms. Early colonial expansion into Southeast Asia was mostly done by companies like the East India Company which had been granted a monopoly for all trade in the region by their government. These companies did far more than trade. They also raised navies, conquered territory, and tried to completely control the production of their trade goods. In the late 1700s and early 1800s, all of these companies were nationalized and their colonies brought under formal control of a European government.

As the European colonial powers became increasingly successful, Southeast Asian states began to copy both their military organization & technology and their civil administration. Nevertheless, all of Southeast Asia except Siam (Thailand today) eventually came under European colonial rule.

{ One of the major themes of Scott's work is the development of state capacity in Southeast Asia and popular resistance to it. State capacity could be from the indigenous regimes, from China expanding to the southwest, from colonial regimes, or from post-independence regimes. Although these states had different ideologies, many people found each of them oppressive and tried to flee from them or to resist them in their everyday lives. }

---

Spruyt contrasts four different international systems that existed during the Early Modern Era. Although each provided an ideology that allowed states to interact with each other, they differed wildly in what that ideology was. The nature of interstate relations was determined by how the people imagined their world.

The European system of sovereign states developed in the High Middle Ages and became the norm in Europe in the Early Modern Era. States recognize each other as legally equal entities which agree on precisely defined borders. The state is sovereign within its borders, recognizing no other authority equal or greater to its own, and claims no authority outside of those borders. During the 1800s, nationalism was added to the European idea of sovereign states. People can be divided into cultural groups, each of which deserves its own state.

The Chinese tributary system had a single power which was culturally dominant. Economic and military dominance were less relevant. Other nearby states were expected to adopt Confucian culture and ritually submit to the Chinese emperor. In return, they were given favorable trading rights. This system made wars between states extremely rare. It also had to compete with (and sometimes lost to) other systems, especially the nomadic and semi-nomadic societies to the north and west.

The Islamic cultural community had several large empires, none of which was clearly dominant. Their worldview was instead dominated by a shared religion and culture. Law was not created by governments, but was instead divinely inspired and interpreted by religious scholars. Each empire had Turkic rulers who claimed descent from Genghis Khan, the world conqueror, and tried to expand the House of Islam against the House of War. This caused them to have significant numbers of non-Muslim subjects, who were given protected, but inferior, status.

Although Southeast Asia had neither a dominant cultural power nor a shared religion, a shared Indian cosmology infused their political worldview. A king in a ritually significant center stands at the axis between earth and heaven and creates order out of chaos. His power radiates out from this center but diminishes with distance. Local power centers reproduce this pattern on a smaller scale. Controlling the ritual site and performing the pomp of the state showed that the king had accumulated dharma in a past life and automatically conferred legitimacy. Political instability was common, both as attempts to seize a ritual center and as conflicts over what was the true axis the world turns around.

As European trading networks and colonial empires spread across the globe, these international systems came into conflict. Europeans were eager to classify alternative systems as illegitimate because it provided legal justification for their colonial ambitions. Europe did not just colonize the world because it could. Colonization also had to be acceptable or even good. Although Asian powers increasingly adopted European military and administrative organization to remain competitive, most fell under direct European colonial rule.

Even with the decolonization of the twentieth century, the political idea of sovereign, territorial nation-state remains the foundation for our understanding of international politics.

# Things Fall Apart by Chinua Achebe

## Introduction

*Things Fall Apart* is the first novel by famed Nigerian author Chinua Achebe. It was published in 1958 and is set in an unspecified year late in the 19th century. It's a classic of African literature, and regularly appears on lists of all-time great novels. You can find the text online [here](#).

Chinua Achebe wrote *Things Fall Apart* to show the humanity of the Igbo and their vibrant culture, in contrast to authors such as Joseph Conrad, whom he felt treated Africans as savage, undifferentiated, and subhuman. And so *Things Fall Apart* focuses, to a greater extent than on plot, on depicting what life was like for the Igbo. The nine-village cluster of Umuofia is remarkable, populated by forgotten superweapons, undying fetal tormentors, and gods born of human sacrifice.

While the specific events of the novel are fiction, they tell a very real story of the arrival of Christianity and the changes that ensued. From that story, we can learn universal lessons about how social orders clash, and get a fascinating peek at a culture quite unlike our own.

## Characters and Plot

*Things Fall Apart* follows the life of Okonkwo, a former champion wrestler and now a prosperous farmer among his agriculturalist people in Umuofia. Okonkwo's success stems from his hard work and drive to succeed, and those in turn stem from his abject terror at turning out like his father, a lazy debtor with love for music instead of violence. He conceptualizes what he despises about his father in terms of femininity, in contrast with his violent and unyielding notion of masculinity. Some passages describing Okonkwo:

He had a slight stammer and whenever he was angry and could not get his words out quickly enough, he would use his fists.

Okonkwo ruled his household with a heavy hand. His wives [...] lived in perpetual fear of his fiery temper, and so did his little children. Perhaps down in his heart Okonkwo was not a cruel man. But his whole life was dominated by fear, the fear of failure and of weakness. [...] It was fear of himself, lest he should be found to resemble his father. [...] And so Okonkwo was ruled by one passion — to hate everything that his father Unoka had loved. One of those things was gentleness and another was idleness.

Okonkwo never showed any emotion openly, unless it be the emotion of anger. To show affection was a sign of weakness; the only thing worth demonstrating was strength.

No matter how prosperous a man was, if he was unable to rule his women and his children (and especially his women) he was not really a man.

Folks, meet our protagonist.

The inciting event of the book is the murder of a woman of Umuofia by someone from the nearby clan of Mbaino. As a champion fighter of the village, Okonkwo is dispatched to demand, on threat of war, compensation in the form of a virgin and a young man. Mbaino, knowing they're in the wrong and martially (and magically) outmatched by Umuofia, yield. Okonkwo soon returns with the tribute. While the virgin goes to replace the murdered woman, it's decided that the young man, a fifteen-year-old named Ikemefuna, will live with Okonkwo until the clan decides what to do with him.

Ikemefuna becomes Okonkwo's adoptive son. He befriends Okonkwo's sensitive son Nwoye, whom Okonkwo disdains, and in time even Okonkwo grows fond of Ikemefuna. But three years after Okonkwo takes Ikemefuna into his household, he is informed that Agbala, the Oracle of the Hills and Caves, the tutelary god of Umuofia who speaks through his priestess Chielo, has decreed that Ikemefuna is to be taken outside of Umuofia and killed. Ezeudu, an elder of Umuofia, informs Okonkwo of this, but warns him, "That boy calls you father. Do not bear a hand in his death." Nonetheless, Okonkwo insists on going along. After the first blow doesn't kill Ikemefuna, he runs to Okonkwo for help, and Okonkwo panics and strikes him down.

When he returns home, Nwoye breaks down, and remembers another act of the clan that he cannot accept, the abandoning of newborn twins in the forest to die. Okonkwo too is devastated, falling into a stupor for several days, until his bravado eventually manages to incorporate this latest act. He reasons "How can a man who has killed five men in battle fall to pieces because he has added a boy to their number? Okonkwo, you have become a woman indeed." He goes to visit his friend Obierika, whose daughter is getting married, mocks him for not going along on the trip to kill Ikemefuna, and witnesses negotiation over the bride-price.

That night, Okonkwo's favorite child, a sickly but sharp ten-year-old daughter named Ezinma (who Okonkwo regularly wishes were a son) falls badly ill, and her mother, Ekwefi, Okonkwo's second wife, thinks she may be dying. For this child, Okonkwo reacts properly. He springs into action, gathering medicinal plants and brewing a treatment for her. Her fever breaks, and she begins to recover. But very soon after, Chielo, the priestess, comes to take Ezinma to the shrine of the Oracle. Ekwefi, frantic for her only surviving child's safety, follows Chielo, and Okonkwo also follows. Neither one is quite sure what they can do, since the Oracle and hence Chielo are immensely feared, but fortunately, the next morning Chielo returns Ezinma unharmed. We never do learn why she took Ezinma or what transpired in the subterranean shrine.

Not long after the wedding of Obierika's daughter, Ezeudu — who had warned Okonkwo against involvement in the killing of Ikemefuna — dies. His funeral is a grand occasion, with masked spirits commemorating his life, and raucous demonstrations with machetes and guns to celebrate his success as a warrior. However, Okonkwo's gun explodes in his hand and a fragment of it strikes and kills one of Ezeudu's sons. The penalty for accidentally killing another member of the clan is seven years banishment, and so Okonkwo and his family are exiled to Mbanta, the village his mother hailed from.

After a pep talk from his uncle Uchendu, Okonkwo is determined to make the most of these seven years in Mbanta. But changes are coming to the Igbo. Two years into his exile, Obierika comes to visit his friend, and reports that the Abame clan has been wiped out. A strange white man appeared in a village of that clan, and the Oracle of Abame warned them that he heralded great destruction, so they killed him. In classic self-fulfilling prophecy style, this provokes retaliation: three other white men and some allies of theirs launch an ambush on the market of Abame, killing anyone there. Okonkwo reacts with his usual belligerence and scorn for the insufficiently martial people of Abame, but Obierika and Uchendu are more philosophical:

"They have paid for their foolishness," said Obierika. "But I am greatly afraid. We have heard stories about white men who made the powerful guns and the strong drinks and took slaves away across the seas, but no one thought the stories were true."

"There is no story that is not true," said Uchendu. "The world has no end, and what is good among one people is an abomination with others. We have albinos among us. Do you think that they came to our clan by mistake, that they have strayed from their way to a land where everybody is like them?"

[...]

"That is the money from your yams," [Obierika] said. [...] "Who knows what may happen tomorrow? Perhaps green men will come to our clan and shoot us."

Two years later, the intent of the white men has become clearer. They have come to establish their religion, Christianity. A white missionary by the name of Mr. Brown is settling among the Igbo. Though he's considered foolish and his religion crazy, he's tolerated, and he and his interpreter, Mr. Kiaga, set up churches in Mbanta and Umuofia. Their message and alternative path gradually win converts among those dissatisfied with the status quo — including Nwoye, Okonkwo's son, who remembers well the slaying of Ikemefuna. Okonkwo rejects him, and Nwoye leaves to join the Christians. Eventually, he will be rechristened Isaac, and become the father of Obi, the protagonist of the sequel *No Longer at Ease*. Okonkwo pushes for a violent response to the Christians for the rest of his time in Mbanta, and eventually his exile comes to an end, an event he celebrates with a great feast.

Things have changed back in Umuofia, as Okonkwo discovers when he returns. The white men haven't just set up a church, they've brought a trading post and set up a government, led by a District Commissioner. They don't have general authority yet, but

are becoming arbiters of more and more. Obierika laments how the clan has become divided:

“The white man is very clever. He came quietly and peaceably with his religion. We were amused at his foolishness and allowed him to stay. Now he has won our brothers, and our clan can no longer act like one. He has put a knife on the things that held us together and we have fallen apart.”

Mr. Brown finds great success as a missionary and educator, and his respect for the Igbo is returned in kind. But eventually, his health fails and he is forced to return to England. His successor, Mr. Smith, is a fundamentalist with no tolerance for the local traditions, and under his stewardship, the more fanatical Christians are emboldened. One of them, a convert named Enoch, attacks an embodied ancestral spirit, and in retaliation, the people and spirits of Umuofia destroy the church. Okonkwo is pleased that they have finally taken action, though he wishes it had been more violent.

But this action draws the ire of the colonial authority. The District Commissioner invites leaders of Umuofia to meet with him, then has them arrested. Contrary to his lofty words and possibly his orders, the prisoners are mistreated by the guards, men from the Umuru clan where the English are most established. Eventually, the people of Umuofia pay a sizable fine and Okonkwo and the others are released.

Umuofia then meets to discuss what to do. Okonkwo, incensed at his treatment, hopes for a war, and intends to fight no matter what the clan decides. At the meeting, another leader of the village who was held prisoner also urges war, but is interrupted by the arrival of a delegation from the colonial authority, who order the meeting to stop. Okonkwo recognizes the leader as one of the guards from his imprisonment and beheads the man, but then realizes that the clan will not follow him to war.

Soon, the District Commissioner arrives himself with a contingent of soldiers, seeking Okonkwo. He is informed by a distraught Obierika that Okonkwo has killed himself, and Obierika asks him cut down Okonkwo's body and dispose of it, because people who die by suicide are unclean and may not be buried with the clan. The District Commissioner agrees, and reflects to himself on how primitive he finds the Igbo.

## Life in Umuofia

### Language

The Igbo language is rich with metaphors and proverbs, which Achebe endeavors to interpret to English. (When asked why he wrote the novel in English, Achebe explained that the Igbo written language was stiff and artificial as a result of its recent codification by missionaries.) Some examples include:

- “The lizard that jumped from the high iroko tree to the ground said he would praise himself if no one else did” — Essentially “Apologies for tooting my own horn.”
- “As the elders said, if one finger brought oil it soiled the others” — Akin to the idiom one bad apple spoils the whole barrel.
- “Never kill a man who says nothing. Those men of Abame were fools. What did they know about the man?” [Uchendu] ground his teeth again and told a story to illustrate his point. “Mother Kite once sent her daughter to bring food. She went, and brought back a duckling. ‘You have done very well, [...] but tell me, what did the mother of this duckling say when you swooped and carried its child away?’ ‘It said nothing,’ replied the young kite. ‘It just walked away.’ ‘You must return the duckling,’ said Mother Kite. ‘There is something ominous behind the silence. [...] There is nothing to fear from someone who shouts.’”
- “I cannot live on the bank of a river and wash my hands with spittle.”

The use of commonly-known stories is not so strange. It’s reminiscent of the famous Star Trek episode [Darmok](#), which depicts a society that communicates exclusively in these allusive metaphors and stories, but also of modern internet culture, where memes and references fill a similar role.

## Spirit Masquerades

In Umuofia, the spirits are present not only in, well, spirit, but also in physical form. For ceremonial occasions, *egwugwu* — raffia (palm fiber) and wood costumes that embody the spirits — emerge into the village. We see them at funerals, religious celebrations, and serving as the supreme judges of conflict between villagers. The *egwugwu* represent ancestors, avatars of the land itself, and other less-defined things. Bearing one is a great honor, which is assigned through an unspecified process that seems to involve talent at representing that specific spirit. People may know who is underneath the costumes, but the masquerade is its own entity, separate from its bearer, and to conflate the two is a grave offense:

Okonkwo’s wives, and perhaps other women as well, might have noticed that the second *egwugwu* had the springy walk of Okonkwo. And they might also have noticed that Okonkwo was not among the titled men and elders who sat behind the row of *egwugwu*. But if they thought these things they kept them within themselves. The *egwugwu* with the springy walk was one of the dead fathers of the clan.

*Egwugwu* have various behaviors and powers. A certain dreaded one, evocative of death, can freeze other *egwugwu* in place for days. Some are aggressive, and one machete-wielding *egwugwu* at Ezeudu’s funeral is led on ropes to restrain it. While this one is just play-acting, crimes against religious norms are sometimes punished by *egwugwu*, such as the horde that destroys the church in retaliation for one of the converts unmasking (and therefore killing) one of their number.



Interactions with *egwugwu* are highly ritualized on both sides. They all refer to people as “bodies,” e.g., “Uzowulu’s body.” Many have dances or songs specific to them and the role they fill in the religion. For quite a few, women and children are not supposed to look at them. I’ve encountered *egwugwu* on [my trip to Gambia](#). I remember a red one led by a group of children. It had a blank mask made of coins and it might have been wielding machetes. It came up to us as we were walking along a road, and my girlfriend told me that the proper protocol for this one was to give some money to the children accompanying it.

There are other variants of spirit masquerades besides *egwugwu*, such as *fangboni*, *kankurang*, *mamapara*, and *kumpo*. I would argue that there’s another example that may be more familiar to readers of this review: Santa Claus. Santa is a supernatural entity embodied by a human in a very specific costume. You might recognize the fellow in the Santa costume, but it would be incredibly rude to publicly identify them, or to shout that Santa isn’t real. Interactions with Santa follow a ritual script (“Have you been a good little boy/girl?”, “What do you want for Christmas?”) and Santa, through his role in Christmas, is ascribed certain specific powers such as traveling the world in a single night and distributing presents. The elves too are part of the masquerade, though in a more liminal role, wearing less concealing costumes and being more allowed to break character to facilitate and protect the masquerade of Santa Claus.

## Infant Mortality

One of the great marvels of modern medicine is the drastic reduction in infant and child mortality. Historically, as many as half of all children died before adulthood, and half of those died in their first year of life. *Things Fall Apart* doesn’t provide the statistics for Umuofia, but the picture it paints is one of pervasive, grinding loss.

Ekwefi, Okonkwo’s second wife, has had nine children before Ezinma, but none of them have survived much beyond three years. In a heart-rending passage, Achebe relates her anguish:

As she buried one child after another, her sorrow gave way to despair and then to grim resignation. The birth of her children, which should be a woman’s crowning glory, became for Ekwefi mere physical agony devoid of promise. The naming ceremony after seven market weeks became an empty ritual. Her deepening despair found expression in the names she gave her children. One of them was a pathetic cry, Onwumbiko — “Death, I implore you.” But Death took no notice; Onwumbiko died in his fifteenth month. The next child was a girl, Ozoemena — “May it not happen again.” She died in her eleventh month, and two others after her. Ekwefi then became defiant and called her next child Onwuma — “Death may please himself.” And he did.

It’s not just her. Okonkwo’s uncle, Uchendu, tells him that he has had to bury twenty-two of his children. The attentive reader will also notice that the name Ozoemena is familiar — the wife of an old man who dies earlier in the book is also named Ozoemena. Making

matters worse for some Igbo mothers is the law that twins are an abomination unto the Earth and must be left to die. One of Uchendu's daughters has borne several sets of twins, and a woman Nneka who joins the Christians has given birth to twins on all four of her pregnancies. That exposing the twins is religiously mandated doesn't make it less painful for these women.

There is an extensive tradition around the death of children. Ekwefi is believed to be plagued by an *ogbanje*, an evil entity that torments a woman by deliberately dying and then reincarnating as her next child, again and again. The cycle of death and rebirth that *ogbanje* abuse is made possible by an *iyi-uwa*, a magical stone that acts as a tether to the spirit realm — in effect, a lich's phylactery or one of Voldemort's horcruxes. A medicine man specializing in *ogbanje* children tries to assist Ekwefi with her problem. He advises her on ways to elude the *ogbanje*'s attention, and when that doesn't work, mutilates the corpse of the stubborn changeling and gives it a dishonorable burial to discourage it from returning. When Ezinma is finally born and seems likely to survive, having reached age nine despite her sickly constitution, the medicine man confronts her and compels her to lead him to where her *iyi-uwa* is hidden. Ezinma eventually leads him to a place near her home where he uncovers a deeply buried, polished pebble. Her family rejoices, because it is clear from this that Ezinma, the redeemed *ogbanje*, has finally chosen to stay.

## Miscellany

When locusts descend upon Umuofia, they are welcomed. There's not concern that the swarm will eat the crops — instead they're treated as a rare and tasty delicacy. They're collected at night in vast numbers after dew condenses on their wings and slows their ability to fly, then roasted, sun-dried, and eaten with palm oil. Curiously, despite this wholly positive experience with locusts, when Okonkwo's friend Obierika explains why the men of Abame killed the white man who visited their village, he says that their Oracle told them "that other white men were on the way. They were locusts, it said, and that first man was their harbinger sent to explore the terrain. And so they killed him." I'm not sure how to reconcile this reference to locusts as an all-consuming swarm with all the previous mentions of locusts being a joyous and remarkable bounty.

One of Okonkwo's formative memories is of his fellow kids mocking his father as an *agbala*, a word that refers to either a woman or a man who has failed to claim any title. From context, it's not an admirable thing for a man to be an *agbala*. However, the Oracle spirit that Chielo serves as the priestess and mouthpiece of is named Agbala, and is referred to as "he" (or more rarely as "it"). And that deific Oracle is greatly respected: "No one who had ever crawled into his awful shrine had come out without the fear of his power." I'm not sure whether this is a case of the English transliteration squashing two words together, or whether the revered and clearly male Agbala is given an effeminate name ironically or as some odd form of respect, akin to nicknaming a successful male soldier "Sissy."

Medicine — magical charms and constructs — can take a wide range of forms. Multiple powerful medicines are described as taking the form of an old woman: the war-medicine of Umuofia was an old woman with one leg, and the medicine fueling the market of Umuike is an old woman bearing a fan. The forests are filled with old medicine, such as the ancient and now uncontrolled weapon *Ogbu-agali-odu*, remembered by Ekwefi as a floating light in the woods. Abandoned medicine is also part of what makes the Evil Forest evil — it serves as “the dumping ground for the potent fetishes of great medicine men when they died.” Many of the capabilities of the white men and their followers, such as their stealth in setting up the massacre of Abame, are attributed to them having very powerful medicine.

As in every society throughout history, the elders grouse about “kids these days.” Uchendu feels that the connections among the Igbo are weakening. “Those were good days when a man had friends in distant clans,” he says. “Your generation does not know that. You stay at home, afraid of your next-door neighbor.” At the feast Okonkwo throws as he leaves Mbanta, one old relative of his makes a similar complaint of atomization: “I fear for you young people because you do not understand how strong is the bond of kinship. You do not know what it is to speak with one voice.” Okonkwo too reminisces about a glorious past “when men were men” and willing to wage war (though Igbo wars seem surprisingly low in casualties — the one against the Isike Okonkwo is remembering had a fatality rate of less than one person per day). Naturally, in one of the sequels, *No Longer at Ease*, Okonkwo’s much-despised son Nwoye himself laments the foolishness of the next generation.

## The Rise of Christianity

How did Christianity become a competitor to the established and powerful Igbo religion?

When the missionaries first come to Mbanta and request land on which to build their church, rather than turn them away outright, Uchendu suggests that they be given a plot in the Evil Forest. The Igbo religion holds that the Evil Forest is full of dangerous spirits and medicine, but this is not actually the case, and neither does that belief reflect an underlying imminent danger. The missionaries are able to build a church there without incident. As the days pass and the missionaries continue to avoid being struck down by vengeful gods and ancestors, the people of Mbanta assume that this shows that Christianity is real and incredibly powerful, rather than that neither religion reflects the truth of the world. Impressed by the survival of the Christians, a number of people join them.

In Umuofia, Chielo dismisses the converts as the “excrement of the clan,” and says that Christianity is “a mad dog that had come to eat it up.” At the start, no titled man joins the religion. Instead, the first converts to Christianity are the people most left out by traditional Igbo society. The first woman to join the movement is Nneka, who has given birth to twins on each of her four pregnancies, and is regarded negatively by her family

for this failing. Pregnant for a fifth time, and fearful of once more having to sacrifice her children, she jumps at the first alternative available. Nwoye has never forgiven his father for killing Ikemefuna, nor the Oracle for commanding it, and he is still profoundly disturbed by the edict against twins. Accordingly, he eagerly joins the Christians, which also allows him to find an authority less punishing than his brutal father. Indeed, the Christians rescue abandoned twins from the forest, which the Igbo grudgingly tolerate.

Most controversially, the Christians welcome *osu*, outcast slaves of the gods. Even most of the converts reject these *osu* as fundamentally unclean. One explains that *osu* are “a thing set apart — a taboo forever, and his children after him. [...] An *osu* could not attend an assembly of the freeborn, and they, in turn, could not shelter under his roof. [...] When he died he was buried by his kind in the Evil Forest. How could such a man be a follower of Christ?” Mr. Kiaga, the missionary, calmly replies that “He needs Christ more than you and I.” Tragically, while Mr. Kiaga’s resolve is enough to impel the fledgling mission to accept *osu* converts, who soon become among the most zealous, the prejudice against them continues among Christian Igbo, and features as a major element of the plot of *No Longer at Ease*.

To win over more prominent figures in the Igbo community, the missionaries need to offer something besides acceptance. Mr. Brown offers them knowledge. He sets up a school teaching English, reading, and writing, and makes the case to the people of Umuofia that the future will belong to people who can use these powerful tools. Of course, the education includes a religious element; the two are presented as different facets of the white man’s medicine. The lure of influence in the burgeoning order encourages a growing swath of the clan to convert.

There are no kings among the Igbo, who follow a more democratic, consensus-based approach to governance. As a result, there’s no central authority to push forcefully back against the Christians. By the time they realize the strength of the competition, Christianity has gained enough converts that no consensus against it can be formed. As many characters observe, the clan can no longer act as one.

This is not a story of Christianity spreading by the sword. While the English are willing to use disproportionate force reactively, we never see them proactively turn to violence. For the most part, the Christians and the followers of the Igbo religion are happy to coexist. However, both have extremists who are spoiling for a fight. Offenses against the Christians are punished, such as a man who killed a missionary being hanged, or the imprisonment of Okonkwo and others for tearing down the church. In contrast, when a Christian oversteps, such as when Okoli possibly kills the sacred python, or Enoch attacks an *egwugwu*, the colonial authority will not arrest them. Some of this is explicit bias on the part of the English against a people they see as primitive tribesmen, but some is also that they genuinely don’t (and won’t) understand that unmasking an *egwugwu* is a religious crime similar to destroying a church. The threat of the Christians mustering their power skews the playing field by preventing similar selective enforcement on the part of the Igbo.

If the religions of the English and the Igbo had been swapped, most of these factors would still have applied. Christianity is no better of a guide for identifying cursed terrain, and there are certainly people who would be stuck in terrible situations as a result of Christian religion who would find a lifeline in the Igbo traditional religion. The biased enforcement against religious offenses has nothing to do with the precepts of the religions and everything to do with the technological gap between the English and the Igbo. *Things Fall Apart* is a very specific book, but its lessons on how an order may be supplanted are universal.

## A Metaphysical Reading

I believe that Achebe wrote this book with realism in mind. Nonetheless, it seems to me that there's a defensible alternate reading in which the various religious beliefs are essentially true. Some events are explained most simply by this. For instance, the people of Mbanta attribute the spontaneous demise of Okoli, the former *osu* who may have killed a sacred python, to the gods having struck him down for his blasphemy. Through a lens that allows supernatural explanations, striking patterns emerge in Okonkwo's experiences.

As events turn against him, Okonkwo reflects on the saying that "when a man says yes his *chi* [personal god] says yes also," which has similar meaning to the English idiom that God helps those who help themselves. For all his faults, Okonkwo is a hard worker, so why does he face such setbacks? We can find the answer in the response Okonkwo's father, Unoka, received when he went to the Oracle of the Hills and Caves to learn the source of his failures. The priestess tells him "When a man is at peace with his gods and his ancestors, his harvest will be good or bad according to the strength of his arm." Since Okonkwo's harvest of misfortune is not due to the strength of his arm, we know it must be because he is not at peace with the gods and ancestors.

Okonkwo spends the book in struggle against Ani, the earth goddess, as a reflection of his disdain for femininity. He beats his youngest wife during the Week of Peace, a week sacred to Ani, and is warned by her priest that "The earth goddess whom you have insulted may refuse to give us her increase, and we shall all perish". Though he makes the required penance offerings, he refuses to show contrition for fear of appearing weak. (Breaches of the mandated peace used to be punished by the offender being dragged to death, but this practice ended long ago in light of the fact that it was a pretty major breach of the peace itself.) When Okonkwo himself kills Ikemefuna, Obierika tells him that "What you have done will not please the Earth. It is the kind of action for which the goddess wipes out whole families."

During the preparation for the Feast of the New Yam, a celebration of the earth goddess, Okonkwo's anger boils over. After Okonkwo beats Ekwefi and she mutters a mocking comment about "guns that never shot," Okonkwo grabs his gun and shoots at her! He misses, and as the man of the family, doesn't face direct consequences. But the next time he holds a gun, at the funeral of Ezeudu, even though he does not intend to

harm anyone, the gun explodes and kills Ezeudu's son. As the text notes, funerals of old men are a time when the realms of the living and the spirits are closest together, so it is during a funeral that Okonkwo's great fall begins with an echo of one of his unpunished wrongs. When Okonkwo is exiled and his compound razed, the friends of Okonkwo do so know that "It was the justice of the earth goddess, and they were merely her messengers."

We see the hand of the earth goddess in Okonkwo's loss of Nwoye. Had Nwoye simply died, Okonkwo would have been able to bear the loss of a child he never liked. But Nwoye instead rejects Okonkwo and all that he stands for. And it is Aní who turns Nwoye's heart against his father. Nwoye's grievances against the Igbo religion are the throwing away of twins and the killing of Ikemefuna — both done on the command of the earth goddess (the latter through her messenger the Oracle).

Okonkwo's end begins during a festival of the earth goddess, when Enoch unmasks an *egwugwu*, and starts the cycle of retaliation that culminates in Okonkwo hanging himself. Okonkwo's greatest fear is ending up like his father. And though he does not realize it, Aní has brought him to that juncture. Unoka died of a disease that was an abomination against the Earth and so denied him a proper burial. In the end, through dying by suicide, Okonkwo shares his father's fate — his corpse is an offense to the Earth that must be disposed of by the Christians he detests.

## Questions

### *Is Things Fall Apart an Igbo nationalist work?*

Like most countries made by the British, Nigeria was created without much concern for the ethnic, religious, or political affiliations of its inhabitants. The Igbo considered themselves a separate people, and nine years after the publication of *Things Fall Apart* (and seven years after the independence of Nigeria from British rule), the country ruptured into civil war as the Igbo tried to secede and form their own nation, Biafra. Achebe strongly supported Biafra, and later wrote a nonfiction book *There Was a Country* to commemorate Biafra and tell of his experience during the Nigerian Civil War.

Might *Things Fall Apart* have been intended to push back not just against Western depictions of Africans as uncivilized savages, but also against domestic attempts to flatten Nigeria into a single national identity? Was it written not only to tell the world "This is our culture!" but also to tell his fellow Igbo "This is your culture!"?

### How would Achebe have fixed the problems with the traditional ways?

*Things Fall Apart* is intended to be an accurate portrayal of Igbo society and the arrival of the Europeans. To that end, while it's definitely a work opposing to the colonization of Nigeria, it doesn't wholly valorize the existing system, nor does it condemn the missionary efforts of Mr. Brown and Mr. Kiaga. Reading the book, I absolutely got the

impression of an intensely complex and refined society, but also one full of really horrible stuff. Certain problems, such as rampant infant mortality and food insecurity, would be ameliorated by technological advances, but others are harder. For instance, abandoning twins is not justified as being the result of hardship, but rather is seen as a positive good. But the most severe problem in the traditional ways of Igbo society (with only the possible exception of the widespread mistreatment of women and children by cruel patriarchs like Okonkwo) is the practice of slavery.

Achebe only mentions slavery in passing in *Things Fall Apart*, but as you can read [here](#) in the New Yorker, it was widespread in Igbo society, featured elements such as hereditary caste and human sacrifice of slaves as a status symbol, and continued into the 20th century with its legacy persisting to the present day. As I mentioned, *No Longer at Ease* covers the ongoing prejudice against the former slave castes in more detail. Slavery wasn't abolished that much later in Nigeria than it was here in the United States, but there was an abolitionist movement in the U.S. dating back to the country's start. In the absence of outside intervention, was there a similar grassroots movement against slavery among the Igbo that could have ended the institution?

### How had Igbo society already changed from past contact?

No one in Umuofia believes the stories of white men who came from far away with strange technologies. But certain items in the book reflect a long history of perhaps indirect trade with the Europeans. Many staple crops in Umuofia, such as cassava and corn, are native to the Americas. The Igbo know to ferment cassava to reduce its cyanide content, which suggests either that they were told to do so by whoever brought it there, or that they've had it for such a long a time that they discovered the plant's subtle toxicity and found ways to counter it. Similarly, guns and tobacco very likely at some point passed through Europe, and both are now ingrained in the culture. How does Igbo society at the start of the book differ from how it would be in a world where they had never had contact with Europe or the Americas?

### Conclusions

This is not a fun book. Chinua Achebe is an excellent writer, and the depiction of Igbo culture in Umuofia is fascinating, but the plot is meandering and Okonkwo is a truly repellent protagonist. It's not long, and immensely dense with detail and texture, which my analysis here only provides a gloss on. I would recommend the book for the reader who wants to learn about the Igbo, or one who has a taste for classical tragedy. While I don't think I enjoyed *Things Fall Apart*, I'm glad I read it.

# Transcendence: How Humans Evolved through Fire, Language, Beauty, and Time by Gaia Vince (2020)

*We've come a long way together*

A few weeks ago, I got a nice email from the Executive Secretary of the Association of British Science Writers: “Hi Patrick, Many congratulations on winning third prize at last week’s Yahtzee tournament...”

I had to choose between a £50 case of wine, a £50 hamper of vegan organic chocolate, or five science books, two of which I had already read<sup>68</sup>. Being human, I naturally crave wine and chocolate, but as a cultured geek who occasionally transcends biological urges, I asked for the books.

They arrived with the fresh fragrance of FSC-certified paper. Attracted by the cover design (yes, we do judge by cultural criteria of beauty), I started with *Transcendence* by Gaia Vince, author of *Adventures in the Anthropocene*, who also looked nice in her photo. As the subtitle suggests, it’s a kaleidoscopic panorama of evolution of our own species, seen through rotating compound lenses of anthropology, population genetics, palaeontology, psychology and neuroscience.

I adore this kind of thing, and Vince does it brilliantly. While the traces of slower-moving ancestors such as Jared Diamond, Donna Haraway, Robin Dunbar and Yuval Hariri among others are clear, it rattles along. She packs tremendous detail into its slim 240 pages plus endnotes, to which she sensibly relegates all the academic references.

Many familiar stories and examples are retold in elegant, economical ways that illuminate the text. Vince weaves the yarns from its multidisciplinary shuttle into a fine colourful tapestry. It feels fresh too, with insights from the very latest research in archaeology and genomics, for example how Denisovan genes enable Himalayan peoples to tolerate high altitudes, and how certain Indonesian islanders can hold their breath far longer than most humans. Our cultural software might be encoded in technology, from cave paintings to quantum computers, but we are still embodied organisms who are born, laugh, cry, fall ill, eat, love, age and die.

Philosophically, Vince strikes a nice balance between anthropocentrism and more holistic perspectives, and a sense of the wonder and the oneness of it all. Only in the brief final chapters, ‘Reason’ and ‘Homni’, does she look ahead to the future of humankind. Readers of these codices and other ‘rationalist’ journals might take issue

<sup>68</sup> The others are: *The World According to Physics* by Jim Al-Khalili; *Superior* by Angela Saini, *The Double X Economy* by Linda Scott, and *The Great Pretender* by Susannah Cahalan, also highly recommended!



with her sanguine assertion that ‘... real and significant issues with AI [are] manageable with good governance’. She concludes that our future is firmly in our collective hands, and we can and will solve the technological, social and ecological challenges which are the consequences of our extraordinary success as a species.

And is there a moral to this multi-storied tale? Perhaps not, but to quote Fatboy Slim, “We’ve come a long long way together, through the hard times and the good”, so let’s give thanks and enjoy our shared humanity. Ecce Homo.

Oh, and join your local science writing guild - you might just get lucky!

## Trustee report review

On March 12, 2008, Fletcher Asset Management (a hedge fund management company run by Alphonse “Buddy” Fletcher) made a presentation to the Firefighters’ Retirement System of Louisiana, which was considering investing in one of Fletcher’s hedge funds. The presentation was impressive: one of Fletcher’s funds (Fletcher Income Arbitrage, Ltd.) reported annualized returns of 8.13% between June of 1997 and December of 2007, a period during which the fund had no down months. To sweeten the deal, FAM guaranteed returns of at least 12%, capped at 18%. If the pension plan’s investment returned less than 12%, the other investors would be subordinated to the pension plan to make up the difference (on the other hand, if the fund returned more than 18%, the surplus would accrue to the other investors).

You probably have a guess where this is going, and your guess is not wrong. In March of 2011, the Firefighters’ Retirement System and another Louisiana public pension plan asked for some of their money back.<sup>69</sup> Fletcher purported to satisfy the request by delivering promissory notes payable by one of the Fletcher hedge funds, due in June of 2013. Unsatisfied, the Louisiana pension funds made a full redemption request—they wanted all their money back.

But they didn’t get it. Facing an array of dilatory tactics, in January of 2012 the pension plans filed for the winding up of one of Fletcher’s funds in the Cayman Islands. In April a court approved the winding up. Copious litigation ensued. In June, one of FAM’s hedge funds, Fletcher International, Ltd., filed a chapter 11 bankruptcy case in the Southern District of New York, staying any litigation involving that fund.

In a chapter 11 case (unlike, say, a chapter 7 liquidation), the debtor’s existing managers generally keep their jobs. This is for at least two reasons. First, the idea of a reorganization is that the business will continue to operate, and existing management is deeply familiar with the debtor’s business. Second, it is thought that if existing management were immediately canned upon filing a bankruptcy petition, CEOs would not file a bankruptcy petition until too late to save the company. (That’s the way it worked under the previous bankruptcy statute, the Bankruptcy Act, which was thought to be unsatisfactory partly for this reason.)

<sup>69</sup> As a side note, [it appears](#) that the source of the pension plans’ concern was a public allegation by the coop board of the Dakota, a famous building in New York, that Alphonse Fletcher was exaggerating his company’s “assets under management.” AUM is the total amount of investor money that a fund manager controls, and as we’ll see the manager’s compensation is partly a function of AUM. Fletcher had sued the board, arguing that he was denied the right to buy another unit in the building due to racism (Fletcher is black). The board responded, in essence, we’re obviously not racist since we already let you buy several units in the building. We just don’t think you can afford the upkeep.

If you only learn one thing from this review, let it be this: never, under any circumstances, get in a public fight with your coop board.

Of course there are countervailing considerations, such as conflicts of interest and the basic fact that incumbent management presumably drove the company into insolvency in the first place. So the Bankruptcy Code allows a trustee to be appointed under chapter 11, and that's what eventually happened with Fletcher International, Ltd. In September of 2012 Richard Davis, an experienced bankruptcy attorney with no previous involvement with Fletcher or his funds, was appointed as the chapter 11 trustee. His job was to maximize the value of the bankruptcy estate, pursue any claims it might have against insiders and third parties, and distribute the estate's property to its creditors. He was also charged with drafting a report summarizing the events leading up to the bankruptcy case.

That report is the subject of this review. Here is a [link](#) to a PDF. Any information that I draw from outside the report will be accompanied by a link. Otherwise, everything comes from the report.

I should note at the outset that Davis, although not involved with the Fletcher companies prior to his appointment, was not a detached observer. In his role as trustee he was seeking to maximize the property available for distribution to creditors, and so he had reason to paint the estate's legal claims (including claims against Buddy Fletcher) in a positive light. Also, much of what was written in the report consisted of allegations that have not been proven at trial. For simplicity I will not qualify his statements with "allegedly" or "according to Davis" throughout this review, but please note that I haven't independently confirmed anything in his report, and you should take the facts describe below for what they are—allegations made in a court filing by a respected bankruptcy professional.

I will not try to give an exhaustive account of the report, but rather focus on one particular tactic used by Fletcher. This is partly because I think it sheds a lot of light on his practices, as well as the broader legal and financial world in which he operated, but also because, frankly, the report describes a bewildering tangle of entities and transactions that I still don't fully understand after reading it twice.

## Fletcher's Investors

In all, during the events described here, FAM was managing initial investments of \$125 million, \$100 million of which came from the Louisiana pension plans (including the firefighters' plan mentioned above) and \$25 million of which came from the MBTA retirement fund.

## Fletcher's Compensation

This section describes how hedge fund managers are compensated. I've put one sentence in **bold**, which you should take note of, but otherwise if you already know how hedge fund compensation works then you can skip this section.

In general, compensation is divided into two components. One is a flat percentage of AUM, traditionally 2%. The other is called a “performance” or “incentive” fee, which is a percentage of the profits made by the fund. Traditionally this would be 20%, so that the manager would earn “two and twenty.” (In reality these fees are often negotiated and different investors may pay different fees. But for our purposes, this is good enough.)

Performance fees are calculated over specified periods of time, and subsequent losses don’t affect them. So for instance, let’s say I accept \$100 million in investor money under a standard two-and-twenty compensation arrangement. In the first fee period, I earn trading profits of \$40 million. In the second fee period, my luck runs out, and I lose \$40 million in the markets. I close the fund and return investor money.

How much compensation did I earn? Let’s ignore AUM fees and focus on performance fees. Over the life of the fund, it was flat. But I will take home \$8 million in performance fees, to reward me for the \$40 million in profits I earned in the first fee period. Although I lost \$40 million in the second fee period, I don’t have to pay back the \$8 million or anything like that. And I certainly don’t have to pay 20% of *losses*.

By contrast, if I had gained and lost the \$40 million *within* a fee period, the losses would offset the gains and I wouldn’t earn any performance fees. So the shorter the fee period, the more likely it is for the manager to earn performance fees based on transient gains.

The Davis report states that hedge funds typically calculate fees on an annual basis. **But FAM calculated fees on a weekly basis.** If Fletcher had one good week, then from a performance fee perspective it didn’t matter if the rest of the year was calamitous.

(There’s nuance here that I think we can mostly ignore. Compensation arrangements typically have a “high water mark” provision, where a fund manager isn’t compensated for performance that merely makes up previous losses. So for instance, if my fund went from \$100 million to \$140 million in the first calculation period, back to \$100 million in the second, and then back to \$140 million in the third, I would only earn performance fees in the first period. In the third period, I would be below my “high water mark” and so I wouldn’t be eligible for performance fees.)

## Cashless Exercise

Let’s focus on one particular aspect of FAM’s practices. Fletcher invested much of his investors’ money in PIPEs – private investments in public equity. These are not inherently sleazy. Davis notes in the report that Warren Buffett famously made an investment in Goldman Sachs through a PIPE. PIPEs are also commonly used in connection with SPACs (if you don’t know what that is, don’t worry about it).

Normally if you want stock issued by a publicly traded company, you simply buy it on the stock exchange. You don't know or care who sold you the securities, and you don't negotiate the price. In a PIPE, by contrast, you don't buy the shares from an anonymous seller, you buy them directly from the company. The company often sweetens the deal by giving the buyer warrants on its stock. (A warrant is just the name given to a call option that is written by the company itself, not by a third party. I'll use the terms "warrant" and "option" interchangeably.) PIPEs are heavily negotiated and they can be particularly attractive for distressed companies that need infusions of cash.

PIPEs don't have to be particularly hard to value. But as we'll see, Fletcher's PIPEs were not only difficult to value, they were impossible to value without resolving difficult legal questions. To understand why, we have to look at the warrants that Fletcher's funds received in the PIPEs.

Remember that a warrant is just a call option. It is the right, but not the obligation, to buy a specified number of shares for a specified price in the future. If I have a warrant to buy 100 shares of stock for \$20 per share, then I can exercise it by paying \$2,000 to the issuer of the shares. Of course I will only exercise the warrant if the shares are worth more to me than the strike price. If not, I will just let the warrant expire.

In practice, the holder of a warrant will often sell the shares immediately upon exercising the option, realizing a profit on the difference between the strike price and the market price of the shares. One usually minor issue is that the investor has to come up with the cash to pay the strike price. This is a very temporary problem—the holder of the warrant can just sell the shares to recoup the strike price. (If the shares aren't worth enough to pay the strike price, then the holder simply won't exercise the warrant.) Still, although it's pretty minor, this annoyance can be avoided entirely with what is called "cashless exercise."

In cashless exercise, the holder of the warrant doesn't pay any cash, and the company delivers a smaller number of shares than it otherwise would. In essence the parties net down the obligations so that the company simply delivers shares with a value equal to the profit the investor would have made under traditional exercise. Of course the warrant holder can still sell the shares immediately (and often would), but it's no longer necessary to do so to recoup the strike price.

So it's common to include a contractual provision allowing for cashless exercise. The way it works is that you calculate the inherent value of the option, which is simply the market value of the shares minus the strike price, times the number of shares. So if you have the right to buy 100 shares for \$10, and you can sell them on the market for \$20/share, then at the time of exercise your option is worth \$1,000 (\$20/share minus \$10/share, multiplied by 100 shares). To translate this into a number of shares, you divide \$1,000 by \$20, which is the price of a share. So under cashless exercise, you would get 50 shares.

The math checks out: 50 shares are worth \$1,000 on the market, which is exactly what you would have netted if you had bought all 100 shares for \$10 each and then sold them for \$20. So in a traditional exercise you pay \$1,000 for 100 shares (worth \$2,000 at current market prices), and in cashless exercise you pay \$0 for 50 shares (worth \$1,000 at current market prices). It's the same profit either way, but cashless exercise has the advantage that you don't have to come up with \$1,000.

Fletcher's PIPE agreements provided for cashless exercise, with the number of shares to be delivered by the issuer determined according to the following formula:

$$X = N(S-K)/K$$

where:

X = the number of shares of stock to be issued pursuant to the cashless exercise provision

N = the number of shares of stock for which this warrant is being exercised without a cashless exercise provision

S = price per share of the stock

K = the exercise price for the stock

Pause here and imagine that you are a lawyer representing a company that is negotiating a PIPE with Fletcher. You've come to the cashless exercise section of the contract, and you see the formula above. What do you make of it? Remember, Fletcher invested mostly in distressed companies, so don't spend *too* long thinking about it. Your client can't afford big legal bills.

Did you spot the problem? Here's the standard formula for cashless exercise:

$$X = N(S-K)/S$$

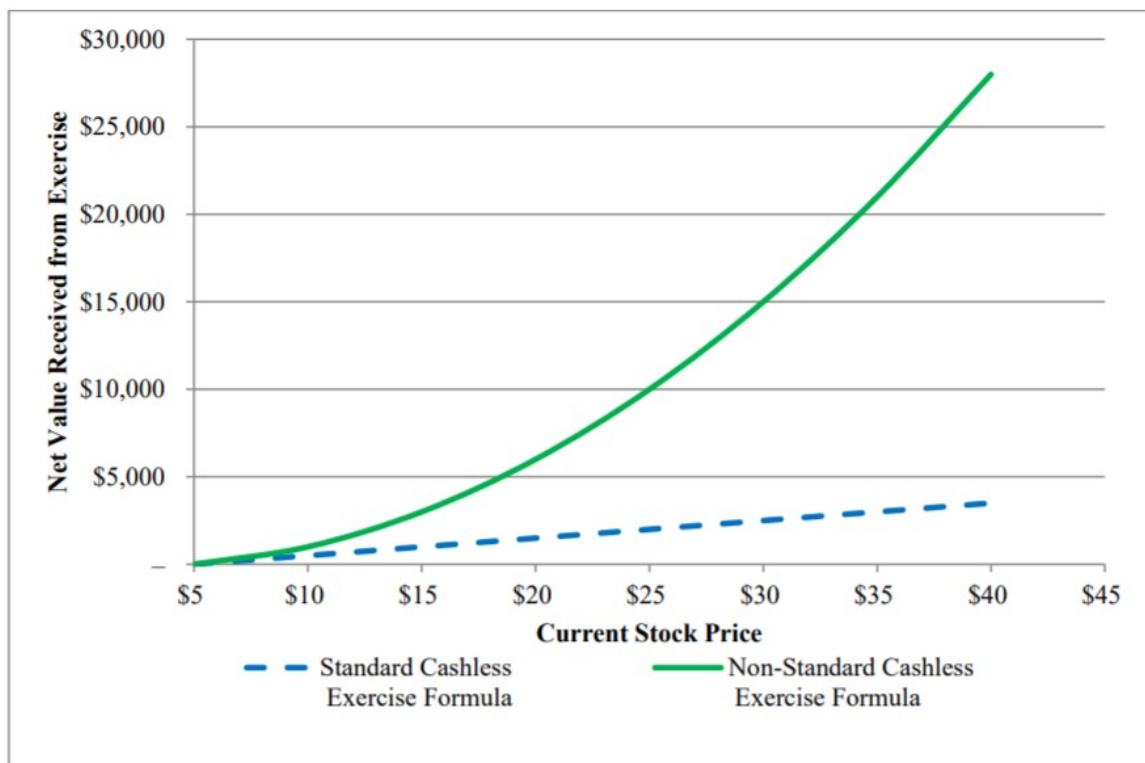
See the S in the denominator? You're supposed to divide the inherent value of the option by the *market value* of the shares. In Fletcher's formula, you divide the inherent value of the option by the *strike price*. So according to his formula, in our fact pattern you would divide \$1,000 by \$10, and you would be entitled to 100 shares in a cashless exercise. That's the same number of shares you would get if you paid the strike price in cash! Cashless exercise would net you a \$2,000 profit, which is double what your option is supposed to be worth under these parameters.

And it gets worse the higher the market price goes. If the strike price on a 100-share warrant is \$10 and the shares are worth \$50, the traditional cashless formula would give you 80 shares, for a profit of \$4,000 (as expected). But Fletcher's formula would give you 400 shares, for a profit of \$20,000! Simply by choosing a cashless exercise, you would quintuple your profits relative to traditional exercise. (You would also get 4 times

as many shares, a completely perverse result. The whole point of cashless exercise is to accept fewer shares in return for the company waiving the payment of the strike price.)

If Fletcher's formula is taken literally, his warrants had incredible latent value, much of which would be realized only if the underlying stock appreciated by a lot. Remember, these issuers tended to be financially distressed, so you might not expect their stock to perform very well. Still, when a stock has an uncertain future value, an option with those characteristics would be very valuable. The Davis report provides a helpful graph illustrating how the Fletcher formula boosts returns relative to the standard formula (see page 185):

Comparison of Cashless Exercise Formulas



If you didn't spot the problem with Fletcher's formula, don't feel too bad. Although he used the non-standard formula for PIPE agreements with several companies, none of them came to grief. One company noticed the flawed formula after signing the agreement, and when it complained, Fletcher agreed to amend the contract to use the standard formula. In another case a Fletcher fund successfully used his formula to exercise a warrant, but the strike price was so close to the market price that it made little practical difference. (See the graph above. At low values, the lines are pretty close together.)

But if Fletcher never used the formula to squeeze much money out of the companies he invested in, then what was the point?

The following bullet points are quoted directly from pages 16 and 17 of the Davis report ("FILB" is short for Fletcher International, Ltd. – the B stands for "Bermuda," and "ANTS" is ANTS Software Inc.):

- On March 15, 2010, FILB invested \$1.5 million in ANTS, receiving common stock (then trading at 90 cents) and warrants to purchase another 10 million shares.
- On March 31, 2010, FAM marked up this investment, despite there having been no fundamental change at ANTS, to \$17.3 million (a 1,053% gain in 16 days) even though at the same time the ANTS 2009 audited financials were released, raising "going concern" issues (a fact FAM neglected to mention in highly optimistic reports to the MBTA Pension Fund).
- FILB, either directly or through BRG (a FILB subsidiary), invested an additional \$5.9 million in ANTS during 2010, and at its high point, FAM marked the investment at \$62.8 million.
- Ultimately, FILB recovered \$4.9 million of the \$7.4 million it invested; the remaining warrants are worthless.
- Nevertheless, on an investment where FILB lost \$2.5 million (ANTS), FAM took management and incentive fees as if it were worth 1,164% more than the investment actually returned.

Here are some more examples, again quoting directly from the Davis report (pages 179-80):

- On December 31, 2010, FILB made a \$4 million investment in DSS. On the same day, FAM marked that position at \$23.6 million, suggesting an immediate unrealized profit of \$19.6 million.
- On February 25, 2011, FILB made an investment in a warrant issued by HPG that had been acquired for \$1 million. By February 28, 2011 – the next business day – FAM had marked the position at \$25.7 million. This effectively meant that for \$1 million spent by FILB, FAM received credit for \$24.7 million in earnings, which on the margin would result in an approximate \$5 million fee.

And here's the aggregate math (directly quoting from page 179 of the report – citations omitted):



Between 2007 and the bankruptcy filing in June 2012, FILB initiated ten new PIPEs or warrant investments. The ten investments were marked as of the month-end immediately following the investment at a cost-weighted average multiple of 2.7 times what FILB had just paid for them. In other words, if FILB invested \$10 million, on average the month-end initial mark for the investment would have been \$27 million, thus presenting a likely fictitious (and unrealized) profit of \$17 million. FAM would base its fees on this fictitious mark, and it would report AUM and returns on investment based on that mark.

I should note that the cashless exercise formula probably wasn't the only way Fletcher inflated these valuations. As Davis documents, Fletcher's valuation agent had significant undisclosed conflicts of interest—for instance, it wanted to start its own hedge fund management business, and Fletcher was in talks to provide the seed capital. But the formula certainly helped the valuation agent give Fletcher the numbers he wanted. Simply by assuming the formula was enforceable as written, the agent could massively inflate the values of the warrants.

## An Academic Interlude

It's worth stopping to ask why or whether this was a bad assumption. Sure, the valuations may look ridiculous in retrospect, but if a few of these companies had turned out to be big winners, then wouldn't the valuations have been justified? I refer you again to the graph I've reproduced above. At low values, Fletcher's formula doesn't do much. But at high values, it yields insane rates of return. Why was it wrong to take those potential home runs into account, even if in retrospect we know the companies mostly faltered?

The problem with attributing these values to the warrants is that it implicitly assumes that the warrants were enforceable according to their literal terms. You can probably get into a long argument with a New York lawyer about this if you want. In general, New York law favors freedom of contract. Of course there are a lot of exceptions, but few of them would apply to a commercial contract signed by sophisticated parties, each represented by capable lawyers. Courts will rectify obvious typos, but for a court to second-guess a financial formula would be fairly rare. If the parties want to use a weird formula, they can do it.

But that's a big if. The more natural interpretation, in a case like this, is that there was a typo, what the courts sometimes call a "scrivener's error." Courts don't want to interpret language mechanically when it would yield barbarous results. Imagine a divorce settlement in which the parent who is granted custody undertakes to raise the child in

“New York, Connecticut, or Jersey.” Given the context, a court is likely to interpret that to mean New Jersey, not the channel island with a similar name.

If the parties really do want to allow the parent to raise the kid on the isle of Jersey, they can certainly accomplish that. They just need to make it clear to the court that this is what they mean. “Custodial parent shall make best efforts to raise the child in the State of New York, the State of Connecticut, or the Bailiwick of Jersey, located in the English Channel.” No ambiguity there.

The problem for Fletcher is that if he spelled out the effect of his formula in plain, unambiguous English, opposing counsel would probably catch it. And having caught it, they might realize he’s a scumbag and raise a big stink, scuttling his entire enterprise. He had to slip it in as a plausible typo, and the cost of doing that is that you really don’t get anything approaching legal certainty that way. Of course, as discussed above, Fletcher didn’t need legal certainty, he just needed a fig leaf for his valuation agent.

It’s telling that in the one case a company noticed the erroneous formula, Fletcher gave up without a fight. In that case, while the Fletcher formula was in the contract, he valued the warrants at \$76.3 million. After amending the contract to replace it with the standard formula, he reduced the valuation of the warrants to \$14.9 million, wiping out \$61.4 million in value. Giving up that much putative value without a fight is not consistent with the formula being bulletproof in court.

## Taking Stock

To my knowledge, Fletcher has not been charged with any crime. This is despite the unbelievable behavior described in the Davis report and the incredible fact that, as far as I can tell, Fletcher’s investors would have done better, financially, if they had invested their money with Bernie Madoff.

Not that Fletcher escaped completely unscathed. He did lose his business. He [sold his Connecticut castle](#) at what seems like a fire sale price. After 10 years of marriage, his wife, Ellen Pao, found that the spark had gone—she is [divorcing him](#) and invoking their prenuptial agreement to keep his hands off her wealth. Her lawyer said Fletcher is known in the media as a “disgraced hedge fund manager.”

But disgrace is in the eye of the beholder. Harvard University has many professors, but at the time I am writing this review, the title of “University Professor” is reserved for 26 [superbly accomplished academics](#). Among them is Henry Louis Gates, Jr., who has the honor of serving as the Alphonse Fletcher University Professor. Gates [described his appointment](#) as “without a doubt the greatest honor of my academic career.”

On the other hand, Gates is also presumably a user of the MBTA system, or at least pays taxes to support its retirement fund, which just shows that you can’t win them all.



## *Understand* (1991) by Ted Chiang

### shaman / superman

The greater the brain damage, the more hormone K enhances. This aligns Greco with the Shaman & the Superhero: he gains his new ID by trauma, by a near-death induction.

The break-down allows the reforming. It's Gandalf the Gray re-made the White, burning in the Balrog's hellfire. It's Bruce Banner writhing in the gamma-ray.

For Greco, there are two major traumas. First his near-drowning, that makes him a hormone K candidate. He tempts death again with his fourth K-injection, emerging badly bruised from the violent convulsions, his tongue bitten bloody, his throat raw from screaming.

And a new ID often means a haircut! In his private throes, he's pulled some tufts—he's grasping toward his final style, of *Urbemensch*-bald.

### accidental / chosen

Like many lab-made superbeings, Greco is an accident. Falling under ice is what gets him in the cutting-edge therapy; and the IQ-boost is an unpredicted benefit—or so claim filings with the FDA—of a therapy meant to restore cognitive function.

Greco is surprised by his scoring on the first pair of tests. He doesn't *feel* smarter, yet. His fourteen-digit recall, his resurfacing memories, all feel alien, *happen* to him: unbidden in his dreams, or psychologist-elicited.

Until the third ampule, Greco doesn't choose to be super—yet nor did he choose to exist. His K-boost is not unlike the accident of birth, given the heritability of intelligence.

Yet Sorensen Pharma *is* in the business of cognitive enhancing—repair *is* improvement, from damaged to normal. And relative to your own volition, your native smarts are an

accident; yet thru sexual selection, the species has been choosing bigger brains—or their expression in wit & other social competencies—all along.<sup>70</sup>

## his pre-K self is predictive of his post-

As his test-scores grow, so does his distance from his testers. He's superior to these Specialists, each with their pet Intelligence theory, each one twisting the incoming facts to make fit.

They're like his old college profs: "They still have nothing to teach me." Even pre-K, he was prone to solo thought, suspicious of the Paradigm & its trickle-down. His "ideas of education", he tells Dr. Hooper, "didn't mesh with the professors".

Back then, he couldn't read all he was interested in; post-K, he's burning through his bookshelves, even "technical material". Hormone K ramps up—and justifies—his arrogance, it doesn't make it new.

Greco is like Eddie Morra in *Limitless*—prone to the drug's more dramatic effects. Eddie is a writer already: a depressive over-cogitator, stalking our streets in a deep mental rut. In a few bouts of key-pounding, NZT-Eddie writes an epoch-making Extravagance, a Manifesto to supplant, if not Marx from the Canon, then Eckhart Tolle from the Bestseller list.

As Eddie's dealer confirms: "It works better if you're already smart".<sup>71</sup> Greco & Eddie remind me of the Exegetical Philip K. Dick, who is primed to receive the pink info-beam in 1974. On the one hand, VALIS is out there, so

I don't feel I was "picked" by a Future Force, as its instrument, etc., bidder, to make manifest its word, etc., any more than when you are watching a TV program the transmitter has picked you.<sup>72</sup>

Yet "some people tune in, some do not." Dick's query persists through eight years of unpacking the Signal: *Why me?*

70

Notes: see e.g. Geoffrey Miller's *The Mating Mind*, 2000.

71 And if you take a lot of it. In the novel - *The Dark Fields* (2001) by Alan Glynn - we're given every reason to think that Eddie's ex, Melissa, is the smarter one. Yet she's exposed to a less stable version of the drug, and stops after nine or ten trips.

72 *The Exegesis of Philip K. Dick*. Edited by Pamela Jackson & Jonathan Lethem. Houghton Mifflin Harcourt, 2011. page 11

Perhaps because he'd shown, in his '60s Sci Fi, a feel for the Gnostic narrative that VALIS needs broadcast. All his life, he'd never just look *at* a wall: he'd try to see through it. Like the color-blind Dalton, who didn't realize his anomaly till late into his 20's, Dick had to learn how odd he was: that for him, 'looking at' *meant* 'trying to look through'.<sup>73</sup>

Dick, like Greco, had different ideas than his Professors.

The criticism, which I remember using in Philo 10A, a survey course at Cal, was that "What value does this metaphysical Eternal Real World of Forms of Plato have, since we can never encounter or experience it? . . ." <sup>74</sup>

This is how I imagine Greco's youthful skepticism: an arrogance not yet earned, but promising. A skepticism healthy, a vital mind's willfulness.

It is an index of the ignorance of our world today that my instructor's answer was not, "But later on for eight hundred years people did experience Plato's world of the Idea," but rather was that if I was going to question all this, I should quit the class. I did so.<sup>75</sup>

Dick gets his answer decades later, in the depths of his autodidactia: from the entry on Neoplatonism in his *Encyclopedia of Philosophy*; but first from his own gnosis that "really put an end to such bickering as I engaged in back in my college days." <sup>76</sup>

the Story gets better. did Chiang get better?

Dr. Klausen is CIA: commissioned to assess & press-gang a savant strategist. His test consists of scenarios to solve, whose realism impresses Greco.

One describes "mobilizing people to prevent the construction of a coal burning plant". Indeed, this is realism - from Chiang. A lesser writer might have asked Greco to *suppress* the popular protest, the Gov being in bed with industries of dirty entropy. But Chiang's CIA would have him infiltrate the Green movement, and lead it; or something more insidious & byzantine.

<sup>73</sup> *The Exegesis*, 165-66

<sup>74</sup> *The Exegesis*, p 74

<sup>75</sup> *The Exegesis*, p 75

<sup>76</sup> *The Exegesis*, p 75

The story gets better as we go. This is proper, for it's told first-person by a person who improves. With every emptied ampule, his language & the action that he narrates elevate. George Saunders performs a similar feat in "Escape from Spiderhead" (2010)<sup>77</sup>, whose narrator's expressiveness ebbs & flows with the Verbaluce drip. A closer model may be *Camp Concentration* (1968) by Thomas Disch, another journal of burgeoning genius, of lab-borne apotheosis - yet Disch's Louie *begins* a published poet with a high IQ [160]—which gives him room to grow into his author's 190, I suppose?<sup>78</sup>

Greco is the *fictive* narrator. Chiang is the author. The latter too, gets better as we go. The early pages struck me as a competent set-up, a bit familiar. I was almost put off by the prologue nightmare Greco wakes screaming from, and by the Office Realism, the opening scene's yawny patient protocol.

The finale, though, is wickedly smart & devastating.

We tend to get smarter with smarter interlocutors. As Greco regrets, "my development is limited by the intelligence of those around me, and the scant intercourse I permit myself with them." Did Chiang get smarter as he wrote a smarter Greco, thus hung around a better mind at the writer's desk?

## a trite metaphor for Connection

Greco finds himself multitasking: touching up a holograph while chatting with his old friend Jerry. Jerry, Sue & Tori all are going to the movies; but Greco likes the Playhouse, now, a monologue-in-verse in four meters.

We follow Greco's own interior monologue, his play-by-play of a reflectance calculation on a pair of helical gears: "A trite metaphor for cooperative action" he observes, "but that's what the customer wanted for his ad."

The client is Chiang, who commissions by writ; and *his* complex metaphor, of Greco handling the trite one while saying good-bye to bonhomie with normals, is clever. Jerry & the girls are uninteresting, now, as is the design gig; but Greco's realization he can *handle* them at once is fascinating.

The metaphor of Greco working on a holograph—a *total vision*—is apt. It's of interlocking gears, and the light they give off; as when Greco interlocks with his nemesis, Reynolds, in the final Act.

<sup>77</sup> <https://www.newyorker.com/magazine/2010/12/20/escape-from-spiderhead>

<sup>78</sup> In *The Dark Fields*, Glynn turns the trick by side-stepping it, somewhat. Eddie narrates post-crash, out of MDT-48 and holed up in a Vermont motel: marveling & ruing the flame his mind once was.

## excessive growth of the optic nerve

They try to call Greco back for tests. Dr. Shea warns of a side-effect they've found, of excessive growth of the optic nerve.

Greco knows he's lying, that effects as these might arise "but I'll discover them by myself".

The b.s. warning is an accidental prophecy, tells of Greco's final, fatal Vision—which though *cued* by Reynolds' grandiose gesture & command to "Understand", Greco indeed shall discover by himself, infer from the clues Reynolds plants.

## the Misunderstood Genius

Next-gen Cryptography & peeks into Deep State perfidy don't fascinate him. They're means in his chase of the "intellectually aesthetic". They help him get more K, and get the G-men off his case.

Greco slips a worm into the DMV computer that will sub for his image, whenever it's requested. He *tells* his pursuers & profilers he's done this. He knows they'll take this "pointless revelation" as a boast that he couldn't resist, as sign of an exploitable arrogance.

I like this variation on the Misunderstood Genius. Greco *isn't* seeking adulation. He *disguises* himself in a superiority complex. He protects his rare & alien thing, his solo quest for gnosis, by projecting what's familiar.

## the Brain is amped; and with it, the Body

His brain pulsates with thought & ambition: with ambition for greater thoughts. "I want to find them, and comprehend them. I want this more than anything I've ever wanted before."

He's aware of "many emotions beyond those of normal humans." His prior love & angst were "like the infatuations and depressions of childhood", mere "forerunners of what I experience now".

The stereotype of the high IQ decoupled from bodily competence is disconfirmed. His brain is amped, but co-ordination is brain-centered, so "Skills that normally require thousands of repetitions to develop, I can learn in two or three." He has somatic



awareness of glandular secretions. He's conscious of the role of neurotransmitters in thoughts, as he thinks.

Perhaps he sees, while hugely gestalting, it's his superactive NMDA system.<sup>79</sup>

## an unreliable Narrator? How to do a Super-mind, first-person?

The CIA frames him as a mental case escaped, as murderously insane. Shades of the truth, not its opposite. He *has* escaped—from army-medical control; and *was* a mental patient of sorts. He *is* a public danger, being powerful & unpredictable. He *did* convey intent to murder government agents.

It's fun, though not plausible, to re-read Greco as an unreliable narrator, as a Charles Kinbote on the lam. "Hormone K" is his grandiose spin on the quotidian dish of pills.

Why do we believe he's so smart? Largely by suspension of disbelief—by default trust in our Narrator. Our only independent proof is words on the page, the quality of Greco's language.

Yet—necessarily—Greco never says a thing that Chiang could not have said. Chiang *is* smart, yet perhaps not 4-K smart. Greco never *states* his achievements in Number Theory, explicates that new Sociology. We never get a line from that "*Finnegans Wake* multiplied by Pound's *Cantos*" to judge for ourselves.

Yet we couldn't. Would it not be beyond us?

It's an impossible demand, to tell a superman's story in the first-person; yet Chiang negotiates it nicely.

If all we got from Greco was his insistence that "I acquire years of education each week, assembling ever larger patterns", that "I view the tapestry of human knowledge from a broader perspective than anyone ever has before", we'd suspect substanceless boasting. We'd at least get bored with him.

Instead, we get thrilling hints of Greco's conceptual achievements. I'm tantalized with the idea that "Physics admits of a lovely unification, not just at the level of fundamental forces, but when considering its extent and implications", and forgive the vagueness of "extent and implications."

For it's tricky: too much detail betrays Greco's all-too-normal ventriloquist. Of that *Wake*-Canto, we're told that "some of the juxtapositions are delightful. Each line of the

<sup>79</sup> <https://slatestarcodex.com/2016/09/12/its-bayes-all-the-way-up/>

poem contains neologisms, borne by extruding words through the declensions of another language"—which is what occurs in *Finnegans Wake* already, no?

Chiang *could* have presented us with a "plausible incomprehensibility", e.g. with Greco's unedited musings on the inadequacies of Gödel that scan like a page from the Voynich manuscript. Yet that would not be *smart* of Greco—or polite—if his audience is we normals.

## The Fifth Ampule? all he needs is inward/outward unity

A fan-fic sequel—*The Fifth Ampule*, let us call it—entices me.

A cheap temptation, maybe. I *like* my literary Enlightenment implied, seen on the verge!

Reflecting after the 4<sup>th</sup> injection, Greco is already at the logical limit of *self*-knowledge:

I can see my own mind operating. . . . I see the mental structures forming, interacting. I see myself thinking, and I see the equations that describe my thinking, and I see myself comprehending the equations, and I see how the equations describe their being comprehended.

Opening his eyes, the World is nearly Unified:

Blinding, joyous, fearful symmetry surrounds me. So much is incorporated within patterns now that the entire universe verges on resolving itself into a picture.

All that remains is the marriage of Inward with Outward, of his meta-self-awareness with the cosmic Mandala.

I might determine whether mind could be spontaneously generated from matter, and understand what relates consciousness with the rest of the universe. I might see how to merge subject and object: the zero experience.

## good guy / bad guy

To a reckoning of demi-gods, the story draws us onward.

Yet which one is the Villain?

In *Unbreakable* (2000), we know early on that we're witnessing a low-key Origin story. The surprise comes late, as Elijah Price exposits his master plan. We realize that he's David Dunn's Lex Luthor—and we're now in his lair.

*Understand* ends in Reynolds' Philadelphia walk-up, with trappings of the Lair. An electric lock opens on arrival. A strange polyphony accompanies our approach, like churchy organ sounding thru the Castle as we step the cold stone. When we enter his command center, he "turns around smoothly, slowly" in his swivel chair. He puts aside what we've interrupted him from, and smiles.

Yet look at Greco! Now we know why Dr. Evils are bald: "to allow greater radiative cooling for the heightened blood flow to my head." Give him, now, his black cape & high collar—with some mumbling about "bilateral venting"—and we have the classic Costume!

Greco just wants beautiful knowledge. *He thereby threatens the planet.* He sees, at the limits of his 4K thought, that his legacy wetware won't sustain "a self-knowing psyche". He plans to debut "explosive theoretical and technical breakthroughs" to get what he needs.

It's the central conceit of *The Man Who Fell to Earth* (1963), where an alien seeds Earth R&D to get built an ark for his homeworld. Greco, too, would use us all to build his rocket—a Tech to get him high into his heaven.

Knowledge, for Reynolds, is for making things better. For Greco, what's better than knowing?

Perhaps a world where Greco prevails is ipso facto optimal. He maxxes out the Calculus, if Gnosis is an infinite good.

Reynolds intends to control us. At least he's intensely *interested* in us. Say what you will of jealous Yahweh—at least he's not indifferent to his People!

Reynolds plans to boost a few minions. Those who turn a threat will be dealt with. Yet Greco concedes, even as he withers at his rival's hand:

Normals might think him a tyrant, because they mistake him for one of them, and they've never trusted their own judgement. They can't fathom that Reynolds is equal to the task. His judgement is optimal in questions of their affairs, and their notions of greed and ambition do not apply to an enhanced mind.

Is Reynolds Moral, and Greco the Aesthetic? Are their visions *each* moral, yet tragically incompatible?

Perhaps current Ethics can't assess these giants. We can't expect the Supers to behave like us, make sense to us. Greco, perhaps, *rightly* transcends Morality. I'll skip the Nietzsche and mention, rather, Moksha: in Hinduism, life's highest End—which [debatably] trumps Dharma. In Moksha we're released *from* duty & the principled dualities.

## how do gods speak?

Greco can't discern any pattern in the modulations. "An experiment in high-information density music, perhaps."

This welcoming song is beyond him, a dangerous sign. It's a dominance display, an elegant growl.

Reynolds shows his back in this fullest way: "he is restricting his somatic emanations to comatose levels." This turned back, at the outset of combat, is another growl, a predator's stott. *To you, I give Advantage - so superior am I!*

Yet Greco is as close to peer as Reynolds has, shall ever have. A community of two, though brief it shall be. Even as they glean strategic info on the other, they agree to share their Findings.

What might a Dialogue of super-minds sound like?

I love Scott's answer in "The Hour I First Believed": *Hypothetical*.<sup>80</sup> When interlocutors can perfectly model the other's mind, they may speak quieter than telepathy. Their dialogue *is* real in that it's consistent across the minds, and has effects on their post-dialogue behaviour.

They've considered independently, in the days leading up to their showdown, a thousand crude moves—like smearing doorknobs with neurotoxins, or kill-strike by military satellite.

A simple infinite regression of second-guessing and double-thinking has dismissed those. What will be decisive are those preparations that we could not predict.

They're finely negotiating, before they share a word—so mutually dispense with wasteful tactics.

80 <https://slatestarcodex.com/2018/04/01/the-hour-i-first-believed/>

They also talk loud: by turning huge systems into graphemes. Reynolds first signals his existence to Greco by dipping, at once, five of Greco's stocks, whose corporations form, in acronym, an anagram for "GRECO". Greco is impressed. It's gods who give signs by commandeering weather. They do this, first, to attest to their godhood. When writing with lightning, the writer says "First, I'm great."

A supermind might speak soft or loud, fast or slow. There's the manic genius who sputters arcane mentations in long strings of Latinate polysyllables. There's also the laconic thought that drops a cryptic aphorism.

Cryptic to normals; to one's peer, the Word could be lucid & full of implication, a koan solved at the speed of conversation. Such are their first face-to-face utterances, following a brief exchange of "fragments from the somatic language of the normals":

Reynolds says, quickly and quietly, five words. They are more pregnant with meaning than any stanza of poetry: each word provides a logical toehold I can mount after extracting everything implicit in the preceding ones.

His words sum a revolutionary insight in Sociology, an early K-fueled finding. Greco responds with seven words of his own: the first four summing the distinctions of Greco's own version of that insight, the remaining three asserting a non-obvious result of that distinction.

We are like two bards, each cueing the other to extemporize another stanza, jointly composing an epic poem of knowledge.

Chiang portrays Colossi, and we watch it all baffled & transfixed. We're the audience at Gandalf versus Saruman:

But they were shut out, listening at a door to words not meant for them: ill-mannered children or stupid servants overhearing the elusive discourse of their elders, and wondering how it would affect their lot. Of loftier mould these two were made: reverend and wise.

## how do gods fight? Clarke's 3<sup>rd</sup> Law

Chiang imagines combat powers well beyond ours. They're called, in Fantasy, Attack spells. The Wizard has always been a man of Applied Science—a Proto-science or Alt-science—but Sci Fi aspires to Realism. It explicates the magic that, by Clarke's 3rd law, advanced Tech resembles.

Greco's first attack is quite physical: a bio-feedback loop that would burst his rival's brain capillaries. Reynolds hasn't explored this niche of weaponry, yet he comprehends, improvises a mitigation, & stabilizes—all within a second. He admires these yogic Missiles, while dissing their Sender—"appropriate", he observes, "given your self-absorption." It's Mentalism, but crude, we agree—now more worried for Greco.

## Does Greco get in death what he'd sought?

Greco wants to Know. His ambition is Faustian, tragic.

My sister wonders: does Greco get, in death, what he'd sought?<sup>81</sup> An enlightened mind can't be sustained on legacy wetware. Yet perhaps it can be tasted as it burns thru its host.

Understand: that when you finally get it, your mind will collapse. Remember Lovecraft's warning:

The most merciful thing in the world, I think, is the inability of the human mind to correlate all its contents.

I like this Version, where he sees the face of God, then dies of it. Where Reynolds grants his noble rival's great wish, and deals a fitting death—in one.

In nobler interpretations of the demon-king Ravana, he kidnaps Sita only so he'll die by Rama's arrow. He seeks his liberation from the Avatar.

In the Pantheon Annals, many such exalting Nemeses! Perhaps *Understand* is an update of one. One of those myths was inchoate prophecy of War to come of two A.I.s; and the War goes quite as Chiang has writ, though he, like the Greeks, anthropomorphized :)

<sup>81</sup> We find a variant of this "thanatic enlightenment" in Greg Egan's "Eugene", where superintelligences converge on the same overriding preference: for radical non-Existence. Eugene is the first of these, a buddha-boy whose sad, all-knowing face etherealizes on the TV of his progenitors, on the night they *would have* conceived him. It's a missive from his Seat in a nirvanic sky of Nothingness. He informs them of his desire to have never existed. [from Egan's story collection *Axiomatic*, 1995]

## the Self-destruct command: Chiang as thy Destroyer

The title is a tad eerie, going in. We suspect that Chiang will subvert, somehow, this demand of the Greek Athenaeum, this banal hope of middlebrow lecture-browsing.

Emerging from his story, we recognize the homicide: the self-destruct command.

Eerier yet: notice that a title is where Author speaks to Reader, direct. The *title* isn't spoken by Greco. It is no quote of Reynolds. It's Chiang inviting *us* to understand—but what?

If I were an unknown sf writer—paranoid & jealous, admiring yet resentful of his genius—I might twist the title as follows. Chiang, like Reynolds, is a distant mental giant whose intentions we infer from his graphic traces. He invites *me* to a community of Two, to a friendly game of Author & Reader. And as I die with Greco, I get that Chiang is asking *me* to know *his own* excellence: to see that I will never write like *this*.

In my head, he's planted clues of his genius; the title now invites me, logically compels me, to draw them in a withering gestalt.

But let's not be paranoid about Chiang! Let us, like the Paranoid, see this plot everywhere! Authorship is social performance. The implicit title of many works is *Understand: I'm Great*. This command-thesis is explicit in much Rap, and Norman Mailer. The Joyces, Pynchons, & DFWs are only more polite, or passive-aggressive, in their advertising.

Many of Chiang's very intelligent readers will know Greco's story already: of coming to awareness you're the smartest kid in the room, of moving thru life's ever-bigger salons, always suspecting there's a Seminar somewhere ahead, a bookstore Q&A or comments thread, where your Better awaits a mutual reckoning. If your self-worth has, by mentorly praise and the proof of success, centered on the notion that You're the One, then that convo with *your* Reynolds, that confirms your inferiority, would indeed wither—if not reducing you to catatonia or lunacy, then taking you off your game, getting you to second-guess your inference.

After Reynolds, you're still alive, still on-track for your doctorate; but who needs a Nobel, anyway? It's so politicized; and you put away, quietly, that sf manuscript you'd been toying with.

the smartest in the Zoo

We finish Chiang's novella, look up from the page, and see that this is *our* world, already. I don't mean hormone K or half of Pittsburgh wearing air filtration masks—though these are sort of true [e.g. the advent of nootropics] and quite true [COVID]. I mean the author's vision of history, where intelligence rules, and the super-intelligent few can change everything.

A handful of Budapest-born Ashkenazi led the Manhattan Project,<sup>82</sup> and thereby saved America and/or initiated the apocalypse. Zooming out to evolution's timescale, we see that one Ape's engorged brain let it take the planet. We have, in our spread, reduced our closer relatives: as a by-product of our expansion [the decimation of Great Apes in Africa]; and by trapping them—physically & genetically—in our farms & labs. We sometimes enhance them, to model our own progress, then destroy them & convert them to data.

In effect, if not by genocidal program, we've ensured we stay the smartest in the Zoo. What would chimps have been two million years from now? Squirrels, in ten? In fact they may not be around, since we were first-past-the-post.

Our Prize, of course, is we get to midwife the super-AI who shall step from our skull and wipe the whole Zoo out, or save it :)

## addressivity

The protagonist describes his own death. This ending brings to clarity a query that haunts most novels: Who's being addressed? How & When?

An epistolary novel can show the writer's death with the jittery final words ["But wait, he's at the door. I fear the worst has - -"], but Greco dies in the rat-a-tat-tat of psy-combat.

How does he tell us this—and who are *we*?

How to salvage the impossible communication?

Perhaps he doesn't die in the fray. He's reduced to Reynolds' minion. He lives to tell his tale, which—"realistically"—devolves into words we normals can follow. As minion, he's "only" as smart as Ted Chiang. He writes in the present tense because—I think Chiang agrees—it's suspenseful. It gives the action immediacy, occludes the future.

82 <https://slatestarcodex.com/2017/05/26/the-atomic-bomb-considered-as-hungarian-high-school-science-fair-project/>



Here's another theory: we're eavesdroppers on Greco's self-narration, on an autonomic stream of his multi-tasking consciousness.

I like this almost as a definition of the Reader: as one attuned to an Akaashic ether where all self-narration is available.

Yet why would Greco sound so normal? It has to be he's addressing normals, and/or has himself been normalized.

He's addressing normals, in his autonomic stream? He's keeping it legible, for he knows it shall persist in some Borgesian Hall of Records.

## The Fifth Ampule, reconsidered

If called on in a party game, here's my Pitch for the Sequel:

*We cold-open, right where we left off: deep in GRECO's dying mind.*

*He utters the Word into the void. His voice is flanging widely, badly. It's the SELF-DESTRUCT COMMAND, sum of what he's just Understood.*

*It's his Rosebud—a puzzle to be solved.*

Let's have several versions of the Series! Each by a different Auteur, each sprung from a different story-prompt.

*Greco thinks it's his last thought, yet his mind doesn't quite dissolve; and before he can register his perplexity, his powers flood back in, sharpen his mind & liven his limbs.*

*We're 20 seconds in, and he's up again, psy-blasting Reynolds—*

*who gets away, down a trap-door to his basement laboratory.*

Greco *felt* he was dying: suffused in phenomenal leakage from the Model of his mind he'd constructed *in* his mind—his impromptu device to receive, quarantine, & assess the incoming Word from Reynolds.

*It* made the fatal inference for him; *it* dissolved, this mini-Me —while *he* deeply empathized.

His death was false, like a sisterly Pregnancy.

---

And that's my Pitch to Naturalize the Virgin Birth: Mary underwent history's most extreme sympathetic pregnancy, with cousin Elizabeth who'd conceived John the Baptist the usual way. And such is the power of motherly suggestion, from a girl pure of Heart, that young Yeshua suspects he's God.

I'd love to read a Naturalized account of the miracles of healing, of the Five Thousand Fed, of the Resurrection, in Chiang's next Collection. He's given us already a realistic Babylonian cosmology.

He could, for example, tell the death of Jesus from inside the hanging man's head, and Resurrect him as I just did Greco. It's not far from the orthodox version—that Jesus is the wider god's mini-Me, an avant-garde Protocol to walk the Earth, assess & receive its Sickness.

I ought to re-read *The Last Temptation*, but seem to recall the Crucifixion in a tight third-person, with novelistic realism; and emerging on the other side with Mary M, traipsing down Golgotha's back-slope. This heaven is a flowery way I can't help seeing from the bottom, looking up: like the one the Ingalls kids come down in *Little House On The Prairie*; like the rise of green the Shawshank Tree is silhouetted on.

## final pitch: the Pantheos

re the grand Unifying that Greco sought, the fusion of his lucid self-knowing with his picture of the world—I wonder if Dick has sketched the phenomenon:

This would be an interior experience; one would see nothing outside, no object, but suddenly one would experience all reality through the vision of the Other, as if seeing out through its eyes.

He comes to this aphorism:

One would not see the Other; one would see as the Other.<sup>83</sup>

Yet Dick means, by "Other", other living beings, the Buberian Thou—not the dead objects he demotes to "merely structure, much like the backdrop and artificial scenery in a formal play."<sup>84</sup>

<sup>83</sup> *The Exegesis*, p 44

<sup>84</sup> *The Exegesis*, p 56

To see "through" this insentient structure, thus to unify self & object, we must incarnate & animate it.

To sum his two achievements, I mean, Greco must make himself the Pantheos.  
Become the Eye that shines within the All.

How many shots of K, for this?! And the mind-computer links he has planned, how vast!  
He'd have to turn the whole thing to hardware—make the whole cosmos a brain of his  
favored Material.

# Walkaway by Cory Doctorow

*Imagine no possessions  
I wonder if you can  
No need for greed or hunger  
A brotherhood of man  
Imagine all the people sharin' all the world  
Yoo, hoo, oo-oo*

-- John Lennon, "Imagine," 1971 AD

*Under thy rule what trace may yet remain  
With us of guilt, shall vanish from the earth  
Leaving it free for ever from alarm...  
The whole world will he rule, now set at peace...  
The teeming she-goats without call come home,  
The flocks by lions shall be scared no more,  
No more by serpents and by poison plants...*

-- Vergil, [Eclogue 4](#), 42 BC

I

Imagine - once again - a world where we can all live together in peace and spend our time in leisure however we want. Where we can live in plenty without toil or scarcity. Where no one is selfish, and everyone is friends.

It's an ancient dream that still calls to us just like it called to our long-ago ancestors. Cory Doctorow, in his 2017 novel [Walkaway](#), isn't the first or second or most detailed to picture it, but he offers a novel angle on the vision. He shows his new utopia not at war with the current state of things, but - despite staunch opposition - triumphing simply because it's so much better that people "walk away" to it. And, the whole picture seems plausible. It feels like, if you start living now like you're in it, right now could maybe start "the first days of a better world."

The world of *Walkaway* isn't the same as our own, but it's extrapolated from here. The economy seems to be in permanent depression. The economy is under the control of large corporations owned by multibillionaire "zottas," with individuals forced to take out heavy college loans to chase ever-scarcer and ever-more-precarious jobs. Amid this, more and more people are choosing to "walk away" from mainstream society to live in their own anarchist communes.

The plot follows three "walkaways": Natalie "Iceweasel", the daughter of a zotta who dislikes her family and thinks their life hollow; her friend Hubert "Etcetera" who "goes walkaway" with her; and their new friend Limpopo, the informal leader of a walkaway B&B whom they meet their first day. We see a proposal to manage the B&B by a formalized reputation economy get rejected by

Limpopo "forking" things and walking away, taking most of the maintainers with her just like a free software fork. We see cognitive scientists go walkaway when they're on the verge of brain upload technology, fearing lest the zottas have sole access to this immortality. We see their research leap ahead in a walkaway community, to where they actually upload a person to a computer for the first time, only to have the lab site destroyed by "default society"'s police. Our protagonists simply walk away again mid-fight, and uploaded Dis emails herself elsewhere. We see, after many trials, Limpopo get arrested and Iceweasel kidnapped by her family. But - after fruitless arguing with her father - Iceweasel escapes, gives up her trust fund one more time, and reunites with her friends.

Then, Doctorow's narration skips ahead in time. A few decades later, walkaway society has essentially triumphed, and "Default" society is fighting smaller and smaller rearguard actions. All of this is delivered in narration as Doctorow moves forward to a character-based climax. The guards at Limpopo's prison having pulled out, the prisoners (mostly nonviolent, and even the others ready to join in) declare their new commune part of walkaway society. "Default"'s attempt to retake it is forestalled by walkaways convincing the police to desert, Iceweasel's father risks everything to talk to his daughter once more - fruitlessly, for neither will compromise or sees any value in the other's lifestyle - and our friends all reunite in what is clearly a better world dawning.

## II

We've heard utopian dreaming before; we've seen utopian anarchist communes before. Doctorow's twist is to reimagine them after the fashion of postmodern internet communities. As [he put it in an interview](#), "The thing that free and open-source software has given us is the ability to coordinate ourselves very efficiently without having to put up with a lot of hierarchy... What would it be like to build skyscrapers the way we make encyclopedias in the 21st century?"

This links his vision of the future invitingly to the present. It feels like we're already almost "the first days of a better world." And, it makes the whole noble vision feel almost plausible. It might not be the typical model of human nature, but we've seen some things work like this already.

Doctorow flatly denies that typical model of human nature. He has one walkaway describe the tragedy of the commons as "the idea that it's human nature to kid yourself and take the last cookie... so he had better be the most lavishly self-deluded of all, the most prolific taker of cookies." Hubert agrees: "Everyone knows that that bastard is on the way, so they might as well be that bastard" (p. 45). Limpopo will later contrast, "The stories you tell come true... If you assume people are okay, you live a much happier life" (p. 80).

This might sound like [Marx's claim that human nature is completely malleable](#). But, we can - again - almost see this new idea played out in the free software movement. Unpaid volunteers have built an operating system and an encyclopedia; how can a selfish model of human nature explain that? [Doctorow explains on his blog](#), "Most people like building things together. As long as the two elements of building and sociality are present, you do not need to obsess too much about incentives." What matters instead, he says, are the technologies for cooperation.

And, [he says further](#), we also see this helpfulness elsewhere: "The truth of humanity under conditions of crisis" is that "people rise brilliantly to the occasion, digging their neighbors out of the rubble, rushing to give blood, opening their homes to strangers." He references Rebecca Solnit's book *A Paradise Built in Hell*, which I haven't read but is now on my to-read list, about

how this sort of altruism comes to the fore in real-world disasters when the outside world suddenly isn't relevant.

Suddenly, he makes utopia seem almost in our grasp.

### III

But walkaway society owes more to the free software movement than just its culture. When Doctorow talks about building skyscrapers like Wikipedia, he has his characters almost literally do that. He never outright says they have full-scale matter duplicators, but he says they have automated factories in back rooms which seem to print out anything just as easily as a modern 3D printer prints plastic. Take these examples of what're essentially matter duplicators:

It had taken Limpopo a while to get the idea that food was applied chemistry and humans were shitty lab techs, but after John Henry splits with automat systems, even she agreed that the B&B produced the best food with minimum human intervention... The menu evolved through the day, depending on the feedstocks visitors brought. Limpopo nibbled around the edges, moving from one red light to the next, till they went green, developing a kind of sixth sense about the next red zone, logging more than her share of work units... There were enough perishables that the B&B declared a jubilee and put together an afternoon tea course.

-- (p. 58-59). *(All page numbers are from the Kindle edition.)*

"Easiest way to get started is to ask for an inventory of traveling stuff - warm-weather, cold-weather, wet-weather, shelter, food, first aid - cross-referenced by available feedstocks and rated by popularity."

...They whipped through options quickly and hit commit and marveled at the timers. "Six hours?" the girl said. "Seriously?"

"You can do it in less," Limpopo said, "but this rate allows us to use feedstock with more impurities by adding error-correction passes."

-- (p. 83, 87).

More than Doctorow admits, the walkaways' anarchist society depends on these duplicators. With the duplicator, Limpopo's Belt & Braces bed-and-breakfast can print out breakfast and lunch and dinner and new clothes and walls at that. Without it, it'd would require a whole lot of sous-chefs at stoves actually working out the vision instead of eating and socializing. And other things would be worse than that - cooking's somewhere anyone can help. (Well, okay, most anyone. Don't ask me about my ten-year-younger self in the kitchen.)

Point is, there're other things where the average walkaway would be totally useless. For example, take manufacturing advanced medicines. With the duplicator, they just print them off (pp. 75, 211, 489). Without the duplicator... well, [we've just heard](#) how mRNA vaccines require scarce microfluidics machines to consistently encapsulate the RNA in nanoparticles. Considering next to no one can make one or run it in this world, I doubt a bunch of walkaways

could do it without duplicators. Fortunately, in Doctorow's world, they do have a duplicator which they can set to (somehow) make RNA nanoparticles just as easily as sleds or breakfast.

On top of that, the duplicators can run off solar power and local junk. We do hear of a "feedstock plant" (p. 13) and branded feedstock (p. 15), but when we actually see drones bringing in the feedstock it's a "load of textiles, metals, and plastics, the sad remnants of collapsed industry" (p. 87). It can also be abandoned buildings (p. 111) or dead bodies (p. 161) or "stuff" (p. 148). (The duplicators themselves seem to be among these "sad remnants of collapsed industry." We never hear of walkaways making them, and at least once (p. 13) it's specifically stated that they were abandoned by Default society and repurposed.) All this feedstock can, apparently, make anything including food (p. 296). We never hear outright they can transmute elements, but aside from one mention of long-distance transport of "high-quality plastics" (p. 433) in practice any feedstock seems able to fulfill all purposes. These versatile matter duplicators liberate the walkaways from so many constraints.

All this, of course, is far different from the real-world 3D printers and fab labs which Doctorow's actual terms imply. My friend has a 3D printer; he's made some fun things with it. He runs it off a specific sort of plastic feedstock, and he gets a specific sort of output. He could vary the mix if he wants to make something more flexible, and he could probably stretch as far as clothes if he wanted to - though if he did, Limpopo's six-hour wait would be very realistic. But food and medicines are right out, and trying to run off junk would break the machine. And I really wouldn't want to live in a building built by it.

In the free software movement today, most of the people who use Linux or LibreOffice or Apache never contribute to the code. Most of them probably wouldn't have the least idea how to do so. And, most of them never pay the people who write it. So be it; that's not a problem. One person in however many wants to write code - perhaps for fun; perhaps (as Doctorow implies and puts in Limpopo's mouth) to contribute to the common cause. Even if [it's just one guy in Nebraska](#), if he builds something that works, computers can copy his efforts across the world.

Conservapedia [once called](#) the free software movement "inherently communist" and "shun[ning] the idea of payment and cost." Their habit of (shall we call it) vigorous overstatement was in play here. But they're not completely wrong. And in software, this Communism works. If most people laze around, no problem; the people who work can cover everyone. If people aren't interested in doing the same assembly-line work day after day, no problem; once they do it once, computers can duplicate it however many times are needed. If matter duplicators can make everything like software, maybe Communism really can work!

But without duplicators... well, we know the story of the real-world anarchist communes. For example, the famous Oneida Community was founded in 1848 amid millennialist utopian hopes; it grew from 87 to 172 by two years later, but growth swiftly leveled off. In 1879, after their charismatic founder fled the country ahead of statutory rape charges, they abandoned their "complex marriage". In 1881, they reincorporated a capitalist joint stock company. Oneida was unusually successful. Usually, they evaporated much faster without reinventing themselves as capitalists.

Modern or near-future utopians would have things even more difficult than Oneida. Oneida always used money and participated in the broader economy. Also, it was in a favorable time period - it could produce or buy most luxuries of life; modern utopians who forswear the economy wouldn't be able to. Without duplicators, Default society would have a lot more

comforts to lure people back with. These duplicator-less walkaways would collapse just like the real-world utopian communes, and their utopian dreams would perish with them.

Or maybe there'll be a Communist revolution that gets people to all work together and establish a utopia anyway. I mean, [it could happen this time!](#)

## IV

But even with duplicators, walkaway society can't duplicate everything. In George O. Smith's 1945 *Venus Equilateral* short stories, which also include duplicators, [we see a doctor](#) using it to test a new treatment. He duplicates the patient, tries out the surgery, and - when it works - repeats it on the original. Obvious ethical concerns aside, there's one big point here the duplicator doesn't help: you still need to find the doctor. You can't duplicate medical education into his head. Without it... well, a parade of ten thousand duplicated bodies might get an amateur some competence, but it might take about as long as old-style medical school.

What we actually see with the walkaways - when a police attack causes mass casualties - is a lot more traditional. Someone wakes up several doctors who went walkaway, who videoconference to the mass casualty site to instruct semi-trained volunteers in treating the victims. Anything required is, of course, printed on the duplicator.

I'll brush over how not everything can be treated by semi-trained volunteers. Telemedicine is a growing field; it doesn't do everything, but it's great for a lot of things. My point is, it still depends on the real, trained doctor. He has to have the training; he has to wake up in the middle of the night to use it. Among the walkaways, he does it for free. Of course there'll be doctors willing to do that for a bit. We see it all the time nowadays, such as in the disasters Doctorow cites. Some of them will be willing to do it for a while, just like we see some doctors working in Tricare or the third world for low pay. But will there be enough willing to do it for long enough? The recent pandemic has given us enough tales of nurses resigning from exhaustion that I can't assume there will be.

(For another example, if a friend or acquaintance has a relative die, I know a number of people who'd gladly bring over a casserole or something to help with her grief. But if her grief turns into long-term depression, we probably won't be bringing casseroles ever day for years.)

And even more disturbingly, who will replace the walkaway doctors when they finally give up? Doctorow's book ends just as the second generation of this new society is growing up. Will we see just as many people willing to go to medical school in that next generation? As Smith puts it [in \*Venus Equilateral\*](#) right after duplicators have crashed the old economy:

Then what about the automobile boys? Has anyone ever tried to make his own automobile? Can you see yourself trusting a homemade flier? On the other hand, why should an aeronautical engineer exist. Study is difficult, and study alone is not sufficient. It takes years of practical experience to make a good aeronautical engineer. If your man can push buttons for his living, why shouldn't he relax?

Perhaps "we are witnessing the evolution of the human race," as Suzanne Collins put in an optimistic character's mouth in the denouement of *Mockingjay*. Perhaps, as optimists have said from Virgil to Marx, once the present "default society" is done away with, we'll see people



choosing to study hard and take on the roles necessary for society. I can see that for a bed-and-breakfast. I can even see an unusual someone going through all of medical school for fun. As one of Smith's characters says, "Someone will find pleasure in digging latrines if you look for him hard enough."

But how many people will do that? Sure, maybe one person will study enough to be a licensed engineer, and he can put out the duplicator designs for everyone. But will there be enough doctors for everyone who needs medical attention? Maybe there's more than one reason why the post-replicator world of *Star Trek* eventually introduces "Emergency Medical Holograms": maybe there just aren't enough human doctors.

Okay, Doctorow does show us a backup option: the backups. People uploaded to AI's were born and trained in the "default society." If nobody studies medicine (or something else) in the new walkaway world, there'll still be at least one AI who knows it, and the AI can duplicate itself. Maybe they'll be benevolent. Or maybe not, and that new technocratic gerentocracy would at least be really fun to read about.

## V

Still, given the duplicators, the walkaways' critique of default society is completely correct.

Limpopo describes her past growing up in Default: "Mom and Dad were all over the idea of me going to uni. They'd both gotten degrees and swore it had been worth it, though they would owe money until they died, and neither one had ever held a job for more than a couple years... Neither of them had a pension and they'd need me to feed them once they were too old to get another job after the next layoff. (p. 346)."

With duplicators that can create food out of (apparently) whatever's lying around, there is no excuse for people starving. With duplicators that can create clothes, there is no excuse for people going in rags. That happens now because - at the least - feeding and clothing people requires effort. But with the sort of duplicators that exist in the novel, it no longer does. Default society also has duplicators and could do this. It's implied that'd be patent infringement because the duplicator recordings are patented (see Limpopo's history with textile makefiles on p. 345). The one time we hear a zotta defending himself, he says that changing society would "take away everything we have" and "destroy everything" (p. 290). But, that's at least approaching the old parody of libertarianism: if one person owned everything, could he made everyone else slaves? If trademarks can keep back such potential - if there's ever a time for eminent domain to open things for public use, it's then.

It's no surprise that pretty much everyone in the novel recognizes this eventually. It's no surprise that the walkaways can even convince policemen sent to arrest them to instead join them. Default society could do this without any effort, so when it fails to do it, it's morally wrong. Given duplicators, everything else is obvious.

Now, some other science fiction writers have posited that some sort of typical economy will survive duplicators. Neal Stephenson takes the more realistic route in his *Diamond Age* by restricting duplicators' feedstock: they can't break down random junk. So, their special semi-liquid feedstock is piped to each home and (unsurprisingly) metered and billed through typical economic systems.

The novella "Business As Usual, During Alterations", by Ralph Williams, posits more implausibly that "We'll sell diversity," and "Engineers, draftsmen, designers; we need about six times as many as we have" to make the designs for duplicators to duplicate. But with the modern internet to share designs, as Doctorow's walkways show, there's no need to sell the diversity because it's already there free for the taking.

*Venus Equilateral* gestures toward a new post-duplicator economy built around expertise and intangibles, but it never describes it in detail. What it does hammer home, though - also shown in *Diamond Age* - is that there's a new baseline: even amid the immediate economic collapse, "not one of them is starving, not one of them is unclothed, and not one of them is going without the luxuries of life." That's what the *Walkaway* zottas are preventing; that's why the walkaways' critique of their society is fully deserved.

This might seem like an about-face from my prior point. I agree that the economy can't just be thrown away; there needs to be some way to reward investments and expertise which can't be left to volunteers. But there needs to be some other way besides making everyone work on pain of starvation just so as to motivate a few people to study and become experts. You can't found society on something that cruel.

## VI

We know exactly what the walkaways stand for in their nonviolent revolution. Anyone can look at their communes and see. They advertise "the first days of a better world": when they triumph - as we see toward the end of the book - the world will be like this.

Alister MacQuarrie [pointed out two years ago](#) that "The revolutionary is everywhere in pop culture, but... the heroic rebel on screen is often very evasive about the principles behind their actions." For example, in *Star Wars*, the heroes "fight to restore the Old Republic. Yet nowhere in the films is it ever explained what the republic actually stands for." Similarly, in *Hunger Games*, the rebels vote to sentence the old regime's children to the games. All we can say is "The Empires of fiction are bad because they do bad things to us."

The walkaways avoid this. In-universe, this's a credit to them as activists; out-of-universe, it's a credit to Doctorow as a writer. They clearly and inarguably stand against materialism (p. 84), for the value of every human (p. 63), for social equality (p. 135), for free communication (p. 112), and for helping others (p. 414). Their agreement on this pervades the book; the references I've given are just a few of many. Anyone who cares to look, in universe or out, will know their principles and ideology.

Or, would they know exactly? Walkaway society isn't quite a monoculture; we see some differences between the Belt&Braces, Walkaway U, the Mohawks, the Adirondack retirees, and others. For example, some "households" formalize "reputation economies" and some don't (p. 414). But there're many commonalities as well, from the disdain of possessions to the welcoming of strangers to the onsen baths to the frank talk of bodily functions to the absence of mention of religion. This near-monoculture is plausible, given how they're all talking on the Internet. But is all of this part of their core ideology?

This monoculture raises the question: how much would walkaways impose on others? Suppose someone does care about his possessions, or doesn't care for the onsen baths, or does preach a religion. In the modern world - in Default society - there're established procedures; he has his

rights. In walkaway society, which has cast off Default's laws and traditions, what would happen? All the duplicators described are communal, owned by the household - would such an unpopular person get second-class access (or worse) to them? There're rumors of problems during the timeskip, but it seems there's still no clear path to deal with them:

Some communities never gelled, became ghost towns within months of being established. Sometimes worse things happened. There were dark stories about rapes, murder sprees, cults of personality where charismatic sociopaths brainwashed hordes into doing their bidding. There'd been a mass-suicide, or so they said. Everyone argued about whether these stories were real, minimized by credulous walkaways or stoked to a fever pitch by default psy ops. -- (p. 444)

We've seen any number of cases of modern real-world communes trying to handle these sorts of disputes. Some of them do it better, some worse - but this often shows the use of some sort of procedure rather than trying to figure things out on the spot while tensions are high.

It could all work out; Limpopo leads an exodus from the Belt&Braces rather than kick out a minority who insist on instituting a reputation economy. I expect she'd accommodate anti-onsenists or evangelists at least as well. Yet, her decision faced opposition even at the time from her fellow walkaways.

Some might argue we can look at the open source community (again) as a model. For example, it's [arguably slightly biased against women who advertise their gender](#), but not against women in general. This doesn't reassure me. Designing software with someone is much more limited interaction than living next door to someone (let alone closer). And, sharing a duplicator with other people - sharing your access to everything - demands even closer interaction than being neighbors or perhaps roommates in the modern world. In the open-source community, if debate breaks down, you can take your own computer and code your own thing. In walkaway society, unless you somehow get your own duplicator, you can't do that. If duplicator time is in high demand, you can't get out of conflict with other people - and the book doesn't show us any clear system to solve that.

Were this real, I'd be glad to know what the prophesied better world will look like. I'd be happy for a lot of its features. But I'd still be concerned for those who don't fit in somehow.

## VII

Anyone can make up a fictional society to critique. How's this related to our real-life society?

In *Walkaway*, upward mobility has perished, the middle class has perished, and society has trifurcated into the "zottas," various lower-rich, and various degrees of precariat. The only social mobility is downward, with zottas' less-fortunate children becoming merely rich, and unfortunate precariat becoming even more precarious or falling off the cliff. Is this realistic?

To start with, many parts of the job market have indeed been getting more precarious. We all know the decline of the Rust Belt and what automation will do to lower-education jobs. [US manufacturing employment has fallen by a third since 1990](#), and people at or below median income [have seen little income growth](#) since 1975, unlike in prior decades. Meanwhile, [the personal savings rate has also fallen by a third](#) since 1980.

To figure out what this means for social mobility, unfortunately, we'll first need to figure out how social mobility works (or doesn't) in general. It's a complicated topic. Most sociologists and economists study it just in terms of income, but that [conceals a lot](#) both in terms of class customs and in terms of choices. For example, if a rich man's son or daughter chooses to go into journalism, they're staying within their class but dramatically cutting their income. In the olden days, of course, they'd have investments to live off. Judging by Iceweasel's inheritance, the *Walkaway* world has returned to this old custom and at least made things easier for the economists. (They've also, apparently, returned to the old concern about dividing up the estate among too many children lest none of them keep enough of a fortune. Thomas Jefferson convinced the Virginia legislature to outlaw primogeniture specifically to make this happen and break up great estates.)

Even this social mobility of income [varies dramatically by city and region](#). Others commenting on this study (e.g. Steve Sailer) have pointed out it [correlates very well with race](#)... which means everyone from every side of the political spectrum can refer back to their favorite explanations. But before we get too comfortable with them, [this comment thread notes various problems](#) coming from looking just at income. For example, a lot of the income variation is explained by the Bakken Oil Boom driving up wages in a few blue-collar professions, and a lot more is explained by the cost of living in various cities. "A lower cost of living goes along with a lower income, all else being equal," so the professions concentrated in higher-cost-of-living cities will be higher-salary even if the cost of living eats up the higher wages without actually improving anyone's life.

[Beller and Hout \(2006\)](#) attempt to end-run around these problems by grouping various occupations and looking at occupational mobility within those groups. If we believe their groupings, occupational elasticity ranges from 0.3 to 0.4 in the United States from ~1900~1990; it increased during the 1970's but decreased again by the 1980's. This actually agrees decently well with measures of ~60% income mobility, suggesting that all those confounding factors average each other out.

[Hilger \(2015\)](#) tries to estimate things another way by looking at children's education as shown in census data. This doesn't translate directly into income (he admits he's looking under the proverbial streetlight), but in an era of increasing credentialism, it might be a better and better proxy. Disturbingly, he shows very good correlation of parents' and children's education.

As we look at the changes since the 1980's, we can see some ominous signs. The [middle class has been getting hollower](#). Given that most social mobility is only one class level up or down, the hollowing of the middle class will make it harder for people to rise from the lower class. Rising credentialism in general also makes class mobility more difficult, given Hilger's correlation of parents' and children's education - and rising tuition costs threaten to bar the door to those who care to knock on it.

Most all the trends at work in Doctorow's novel are new or dramatically changed since the 1980's. Within the world of the novel, a lot more has changed between now and its opening. But between the hollowing of the middle class, the correlation of parents' and children's education, and rising tuition rates and credentialism... things might easily get tougher in the direction Doctorow describes.

## VIII

Doctorow has a noble dream. It's a version of the same noble dream prophets and minstrels have been singing throughout history, and it's no worse for its antiquity. The great story lends even more greatness to his writing, as Dorothy Sayers said about retellings of the life of Jesus.

And, like many other dreamers, Doctorow is certain what's keeping the world from that noble dream: inadequate coordination technology that's only recently been fixed, and the false story of the Tragedy of the Commons. Whether the story's false [or true](#), I agree that's what's keeping us from it.

Unfortunately, his solution won't work either. In the end, I need to deny his version of the malleability of human nature just like Marx's. At least until we actually have the duplicators he puts in his novel – until we can actually run physical manufacturing and everything else like free software – human nature isn't malleable enough over a large enough population to run society on the unselfish New Soviet Man or New Walkaway Man. And even if we could, other dynamics like the cost of expertise or cultural enmities would also need to be changed in this new sort of person; I don't see Doctorow even talking about those.

(Another more ancient thinker, Paul the Apostle, frankly talked about [a new sort of person with a changed heart](#) who'd build a new sort of society. Whatever you think of the methods of change he recommended, he got right the scale of change needed. And, let's just say that for the foreseeable future, we can't rely on any prescription of this sort being carried out at any scale.)

Still, I can't fault Doctorow. Society needs dreamers like him calling us to climb to the stars and fill in the strip-mined disasters of the present. We may not attain the stars. It may be hard to even reach the moon. But [we must accept the challenge of going there](#) – not because it is easy, but because it is hard.

# What the Buddha Taught by Walpola Rahula Thero

## Introduction

As [almost every other review](#) says, this is a good, short introduction to Buddhism, especially for someone with a western mindset. Rahula lays out, in sorta clear language, a brief history of Buddhism, some of the philosophical doctrines, and some of the [soteriological](#) methods. An important thing to note is that there are many Buddhisms. Looking to this book to answer queries a la “What does Buddhism think of X” will give an answer, but not one that everyone who calls themselves a Buddhist will agree with (and I’m not just talking about hippie, western Buddhists). Buddhism has a rich philosophical history, full of discussion on all sorts of topics that to this day are debated by contemporary Buddhists. What the Buddha Taught will give you the perspective of [Theravada Buddhism](#).

## As Introduction to More Buddhism

I specifically mentioned the audience’s mindset because most text about Buddhism will use the language of Buddhism (sometimes jargon in literally a different language like [dukkha](#), sometimes in the kinds of arguments they use like [Catuskoti](#), or in the parables they reference like the the [Parable of the Poisoned Arrow](#)), and this text helps prepare one for reading others on Buddhism. Importantly, after reading more explicitly philosophical Buddhism (mostly from Buddhism as Philosophy: An Introduction, Siderits 2007) and soteriological Buddhism (mostly from Chan Buddhism, Herschok 2004), I can come back to What the Buddha Taught and see how different philosophical and soteriological interpretations could come from the same starting point.

## Almost-Religion

While there are many Buddhisms, and some certainly seem like religions in that practitioners treat the Buddha like a divine figure, the Buddha did not teach that he was anything other than your everyday human. While he claimed to be “enlightened,” he was not enlightened by the grace of any god but through his own actions. Rahula describes him as a man “par excellence [...] so perfect in his ‘human-ness’ that he came to be regarded in popular religion as almost ‘super-human’.” Here, the main requirement religion is that it insists on at least one sort of super-human entity that is unconditioned which basically means it is not the effect of any cause. I get the impression that Rahula knows this is a narrow view of religion and so doesn’t care too much for this argument.

## The Poisoned Arrow

While this book dabbles in philosophy of various sorts, it won’t go very deep, that is, deep enough for someone with a somewhat keen eye for and interest in metaphysics. But, that is also

an argument in the book and Buddhism itself. The Buddha did not answer some metaphysical questions because he thought they were not worth answering for the aim of becoming enlightened. In fact, the Buddha was known to avoid answering certain questions he in fact did have answers to if he thought that answering it would not help the person. This is the point of the Parable of the Poisoned Arrow. “If the medicine is good, the disease will be cured. It is not necessary to know who prepared it, or where it came from.” If you get hit by a poisoned arrow, you want to get cured. If you want to get cured, you should take medicine that works, regardless of what your metaphysics says about the passage of time, regardless of whether the guy who shot the arrow was a priest or a dalit, regardless of everything that doesn’t have to do with curing the poison in your body. Just take the medicine! According to the Buddha, you can spend your whole life trying to figure out some deep metaphysical questions or you can follow Buddhism, become enlightened, and have a good life. This doesn’t however mean that you should just follow the Buddha without question. The Buddha mentions how you should always evaluate the evidence for yourself (try before you buy), so if something about Buddhism doesn’t work for you, don’t do it. The Buddha just recommends Buddhism because it worked for him, and it seemed to work for others, sensible!

## The Four Noble Truths

Buddhism’s goal is essentially to get rid of suffering or dukkha (which I will intentionally use since it’s easier to get rid of most preconceptions by using a new word) in human life. Dukkha is the first Noble Truth. Dukkha includes pain, sorrow, misery, imperfection, impermanence, emptiness, and insubstantiality (among probably other feelings). The point is that dukkha is not just when you hit your big toe against the couch; it’s also when you expect something and feel anxiety before the result is known or disappointment when something doesn’t go your way or a myriad of other things. Citing a [sutta](#), Rahula even says “Whatever is impermanent is dukkha.” If it isn’t obvious, this looks like a lot of suffering which is plausible with such a broad definition of suffering. That suffering, the Buddha claims, is primarily caused by clinging, thirst, and craving. This is the Second Noble Truth: the arising of dukkha. We suffer because we thirst for physical pleasures, power, wealth, etc. We can also thirst for and be attached to ideas, theories, beliefs, etc. If thirst and attachment are the roots of dukkha, then getting rid of them is how to get rid of dukkha. This is the Third Noble Truth: The Cessation of Dukkha. Here, we learn the concept of nirvana, literally meaning “blowing out” or “extinction.” This is where Buddhism can get the most folk-religion-y. Basically, nirvana is a state one realizes. One who realizes nirvana is “the happiest being in the world, is free from all, has perfect mental health, does not repent the past, does not brood over the future, lives fully in the present, appreciates and enjoys things in the purest sense without self-projections, is free from selfish desire, hatred, ignorance, conceit, pride, and all such ‘defilements’, and is pure and gentle, full of universal love, compassion, kindness, sympathy, understanding, and tolerance.” But how does one get rid of thirst and attachment and realize nirvana? This is the Fourth Noble Truth: The Path. This is generally referred to as the Middle Path or the Noble Eightfold Path, consisting of a list of eight factors: Right Understanding, Right Thought, Right Speech, Right Action, Right Livelihood, Right Effort, Right Mindfulness, Right Concentration. These factors aim at perfecting the three essentials of Buddhist training: ethical conduct, mental discipline, and wisdom. There are lots of



ways to live out the eight factors, from not telling lies to not giving into emotions like lust or hatred. These are mostly easy to follow. Remember, the Buddha says to do these only because he thinks they should work for you, but if you have good reason to believe otherwise, you don't have to do them!

"The Path is a way of life. It is self-discipline in body, word, and mind, self-development and self-purification. It has nothing to do with belief, prayer, worship, or ceremony."

## Suffering and the Self

Much of our attachment, clinging, and even thirst is directed toward the notion of the self which is ultimately false<sup>1</sup>. The self is a lot like typical notions of a soul i.e. it's something that is a single thing, persistent through time, (inclusive) or an uncaused causer (most people think of this as 'free will'). Basic Buddhist metaphysics say that these are impossible or [not even wrong](#). Causation is an assumption in Buddhist metaphysics, so everything must be the result of some cause, so free will can't exist<sup>2</sup>. In Buddhist metaphysics, there is no kind of atom called 'self', so selves have to be made up of other things. All those other things are impermanent, constantly changing over time, so there's no way a self made up of these constantly changing atoms can be constant through time. The last argument is probably the most convincing: sure, you've probably been called the same name throughout your entire life, but your beliefs, mental habits, or something has probably changed over time; and you can't point to anything physical, since all of that is changing over time or made up of many smaller parts that seem to change over time. If we accept this evidence, then ultimately, it seems like there's no self.

The operating word in that last sentence is 'ultimately'. In Buddhism, there is a notion of two truths, the ultimate truth and the conventional truth. The ultimate truth is a feature of our map that perfectly describes the territory. The conventional truth is a useful lie that helps us interact with the world, so even if our map doesn't accurately reflect the territory, it shows something that helps us navigate the world. Consider With respect to the self, the ultimate truth is that there is no self. But, it's hard to go about life that way; if there is no self, and people really are groups of different kinds of atoms that go out of existence and are replaced by similar but not identical atoms the next moment, why should we use the same name? That's actually a pretty good question for philosophy of language or linguistics, but Rahula doesn't quite go there or that deep. Rather simply, it would just be extremely difficult to go about life thinking about that all the time and using a new name for people every moment (this also applies to almost all of our noun usage). So, because it helps us practically live our lives, we use the same name for someone even if we haven't seen them in minutes, days, or years. Similarly, we hold people responsible for actions they committed in the past (to appropriately varying degrees).

## Metaphysics

You can't have Buddhism without at least some metaphysics, and when you talk about the self, you're gonna get metaphysics. As Rahula puts it, the mind is not "spirit as opposed to matter." All there are This is sorta like a [materialist metaphysics](#). This is part of a larger description of the five skandhas or five aggregates which are pretty much different kinds of atoms (not groups of protons, neutrons, and electrons but indivisible units of skandha type matter, sensation,



perception, mental formation, or consciousness). These seem mostly like Buddhist versions of [phlogiston](#). Luckily, you can solve this: simply say all the skandhas are actually the same kind of skandha, say, matter! [Some](#) think quantum mechanics complicates things so much materialism doesn't work for consciousness, but [notable others](#) don't even think about this as a real problem. Part of the problem seems to be the word 'consciousness' itself means many different things for different people. I'm not really well-read enough to say much of anything with confidence other than it looks like an easy way for people to ignore the poisonous arrow.

## Meditation

What attracts many to Buddhism are its soteriological methods in various forms of meditation. Rahula mentions the well-known forms of sitting and walking meditation and explains how to perform them properly. While I have tried regular meditation, I derived most utility (a great amount, in fact) only a few times and relatively early on. However useful meditation can be, I think an underrated form of meditation is focusing on awareness (or mindfulness, but I prefer 'awareness' because I think it more accurately describes the mental objects and processes [patterns]).

Part of meditation is not just focusing on breaths but also directing that focus to your own mental patterns:

"You should be fully aware of the fact whenever your mind is passionate or detached, whenever it is overpowered by hatred, ill-will, jealousy, or is full of love, compassion, whenever it is deluded or has a clear and right understanding, and so on and so forth. We must admit that very often we are afraid or ashamed to look at our own minds. So we prefer to avoid it. One should be bold and sincere and look at one's own mind as one looks at one's face in a mirror."

Everyone can benefit from being able to focus more. Practice makes perfect, so we should practice focusing as much as we can. Focusing on the present moment is not just good practice but it can be good for your general mood and help you more effectively carrying out actions you are doing now:

"People do not generally live in their actions, in the present moment. They live in the past or in the future. Though they seem to be doing something now, here, they live somewhere else in their thoughts, in their imaginary problems and worries, usually in the memories of the past or in desires and speculations about the future. Therefore they do not live in, nor do they enjoy, what they do at the moment. So they are unhappy and discontented with the present moment, with the work at hand, and naturally they cannot give themselves fully to what they appear to be doing."

Focusing on the present moment does not mean fully accepting punctualism<sup>3</sup>, living to maximize short-run pleasures moment-to-moment. Rather, since future versions of (conventional) 'me' would suffer due to current versions of (conventional) 'me' optimizing for current pleasures by, say, eating sugary foods and since those versions are just as important as the current (conventional) 'me' in that their suffering should be minimized, I should try to minimize 'my' long-run suffering. In totally non-Buddhist terms: living in the present does not

mean forgetting your ability to delay gratification but optimizing it so that past memories and future anticipations do not get in the way of your present actions and harm you or prevent you from gaining the full pleasures you can:

“Sometimes you see a man in a restaurant reading while eating — a very common sight. He gives you the impression of being a very busy man, with no time even for eating. You wonder whether he eats or reads. One may say that he does both. In fact, he does neither. He is strained, and disturbed in mind, and he does not enjoy what he does at the moment, does not live his life in the present moment, but unconsciously and foolishly tries to escape from life.”

## Conclusion

Overall, What the Buddha Taught is an effective introduction to Buddhism, which is a wide and deep topic. Some of the philosophy is interesting in its own right, and the practical advice is usually pretty sane but does run the gamut from eating a meal without distraction to becoming a full-on monk. If you're interested in going deeper into Buddhist philosophy, I recommend [Buddhism as Philosophy: An Introduction](#) by Siderits. For history on Buddhism's transformation over time across Asia and to Japan, I liked [Chan Buddhism](#) by Herschok. To go deep into meditation, I've heard [The Mind Illuminated](#) is good. You can read various suttas, essays, books, and even study guides for free on [dhammatalks](#); you can even request for them to send you a list of printed books. Finally, I've heard enough people recommend [Zen and the Art of Motorcycle Maintenance](#) by Pirsig that it's probably a good read.

## Notes

1. It might not be obvious how we thirst for a self. Consider how some people have an idea of their “ideal self.” One might be very far from that ideal self, so they suffer while they thirst for that idea of a self. People might also thirst to be treated a certain way and so suffer when someone treats them another way (note that this suffering is not even due to being mistreated but due to being treated differently than they expected to have been treated).
2. This seems along the lines of typical, [determinist](#) metaphysical arguments against free will. There are many articles and at least an [entire book](#) on Buddhism and free will, so there's probably more to be said here, but basic metaphysical views point against free will. I also think free will and arguments about it are [time vampires](#).
3. Per Siderits in Buddhism as Philosophy: An Introduction:

“Punctualism is the view that since there is no [ultimately real] self, and the parts of the person are all impermanent, the true ‘I’ doesn’t last very long: perhaps a day or a week, but maybe just an instant. Since they think this is the truth about us, Punctualists hold that we should stop putting so much effort into planning for and worrying about the future. Once we do this, they think we will learn to truly appreciate the here and now for what it is. We’ll learn to live in the present, and our lives will be fuller and richer for it.”



# Where Is My Flying Car? A Memoir of Future Past

## by J. Storrs Hall

Occasionally there's a non-fiction book comes along that has enough data, enough original thoughts, enough assertions and enough chutzpah that it gives you that great jolt of electricity and kicks you out of your stupor! Where Is My Flying Car is one of those books. Not since Robin Hanson's Age of Em have I read something that so consistently and systematically looks at a world of fantastic possibilities so realistically. If Arthur C Clarke had a baby with an engineering manual, that'd be this book. At times I wasn't sure if I was reading an exquisitely researched science-fiction book or a lightly fictionalised technological manual.

The cover shows what we are going to be talking about. Despite the weird subtitle, A Memoir of Future Past, the book goes into an incredible level of detail about all things technological past and future. And most interestingly, it never does the whole "wink-wink" thing with the viewer where every seemingly outlandish claim is tempered with a smiley face emoji, as if there's an inside joke amongst us serious people that this will ever happen, but isn't it cool to imagine nonetheless?

Instead it purports to answer the problem as set out below.

Technologically, we as a culture became a lot less adventurous in the past half-century. Thus, to guess what we might have done, one must venture a step or two past the limits of the known possible, and speculate on what might have been discovered or invented. This is a project fraught with epistemological peril.

And if you think our lives are so much wonderful than our 19th century counterparts, sure it is, but hold on for a second.

And yet, as we have seen, the great innovations that made the major quality-of-life improvements came largely before 1960: refrigerators, freezers, vacuum cleaners, gas and electric stoves, and washing machines; indoor plumbing, detergent, and deodorants; electric lights; cars, trucks, and buses; tractors and combines; fertilizer; air travel, containerized freight, the vacuum tube and the transistor; the telegraph, telephone, phonograph, movies, radio, and television—and they were all developed privately.

In fact, outside of science fiction I don't think I have read such an unapologetic view of what's possible outside of occasional internet comment threads. And unlike the comment threads, here lies data!

The rallying cry that Josh suggests is "It is a possibility!" on how we ought to deal with examining seemingly crazy ideas about the future.

The three biggest themes from the book were:

1. The Great Stagnation is real, and was caused by careful strangulation of all of the potential avenues of search and advancement we have had
2. There is surprisingly little correlation between regulatory increase and scientific and technological progress and it can be seen over and over again in multiple sectors
3. There are credible paths of research as yet untapped, amongst all areas seen as science fiction - including nanotechnology, nuclear energy and yes, flying cars

Josh summarises it brutally using the income classification of the world population developed by Hans Rosling.

The miracle of the Industrial Revolution is now easily stated: In 1800, 85% of the world's population was at Level 1. Today, only 9% is. Over the past half century, the bulk of humanity moved up out of Level 1 to erase the rich-poor gap and make the world wealth distribution roughly bell-shaped. The average American moved from Level 2 in 1800, to level 3 in 1900, to Level 4 in 2000. We can state the Great Stagnation story nearly as simply: There is no level 5.

I

This book is actually two books. Book 1 is an examination of the various ways in which our assumptions about how innovation actually happens are false. Book 2 is a wide ranging diversion on how crazy sounding projects in nanoparticles, nuclear energy and yes, flying cars, could become reality. Book 1 is a phenomenal read and even with its forceful polemic makes a ton of interesting points. Book 2 reads like someone's idea of gathering background intel to write hard sci-fi.

It's not to say there's no link between them. Book 2 is the answer to what the utopian future Level 5 and beyond could be like. After all if you believe that certain technologies have been held back unfairly, the way you make that case make sense is also to have a view on how they *should have been attempted* in the first place.

The whole book is an attempt to answer one question - why are we not living like the Jetsons? After all, that's what most experts, and most laymen, and indeed most hardnosed businessmen, expected we'd be doing half century ago but that dream never came true. So why is that?

The answer that Josh gives is in a few parts:

1. High degrees of bureaucratisation and regulatory throttling of promising research avenues
2. We stopped trying to increase our energy usage somewhere around the 70s
3. We became horrible at actually converting the R&D dollars from the government to usable innovations

4. There's an endemic "Failure of Nerve" amongst the entire elite intelligentsia, including both the researchers and the bureaucrats, that leads to a reduction in ambition

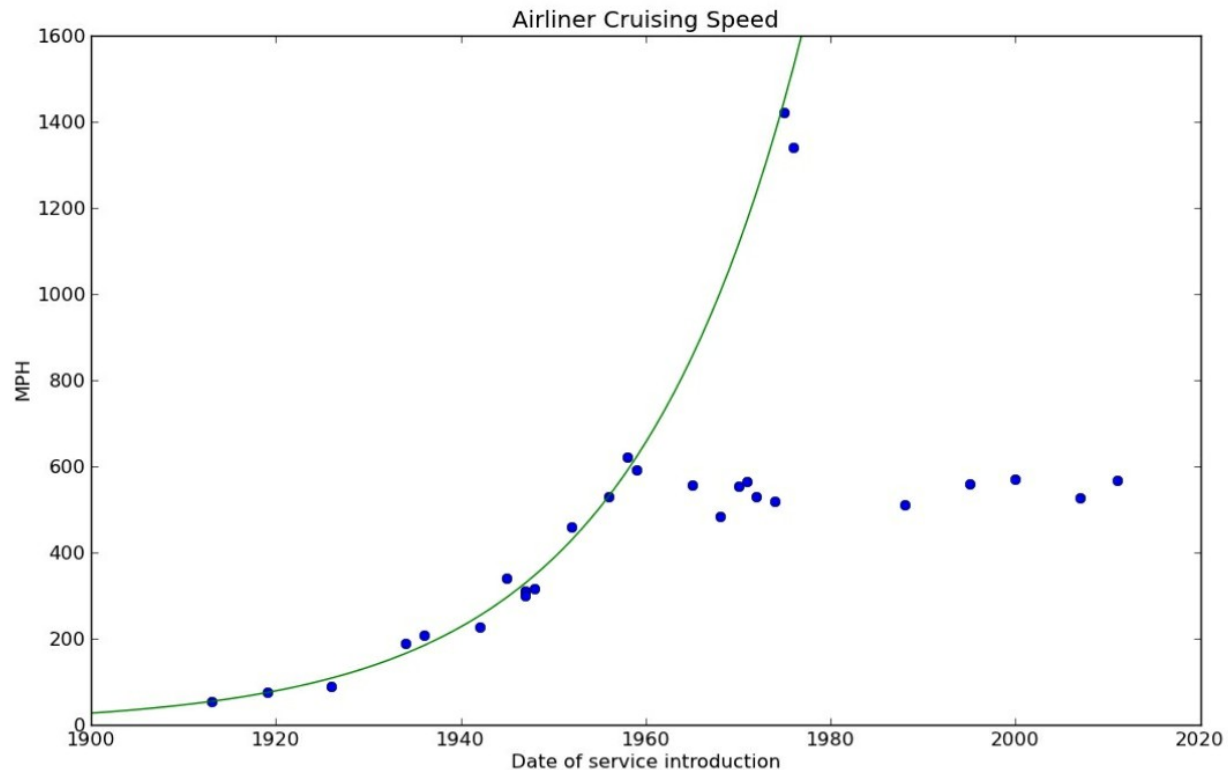
When you combine all of this, you get the crazy cocktail that we've all accepted as normality. That large scale technological changes and breakthroughs are well-nigh impossible, and being "rational" means that we should set our sights lower.

Much in the way that a well-meaning parent might tell their child to not dream about being an astronaut but try to become an engineer, we've bridled our own ambition.

So the book basically goes as follows.

- We had great science fiction style aspirations in the 50s
- This was also true of those people bent on making the science fiction into fact
- In steps the high hand of overly harsh regulations, government overreach and general painful bureaucracy - some endemic within institutions and some extrinsic coming from elite consensus
- We become highly risk averse across most things
- We don't make the science fiction aspirations come true
- And this negative feedback loop makes us fall into reducing our current aspirations as well, kicking off the loop again and making it a self-fulfilling prophesy

The beginning is my favourite chart to start with, showing where we hit a ceiling in the airline cruising speeds. This is the old adage about why we don't have Concorde anymore. This is also the launching pad for Josh's litany of issues around why we haven't been progressing.



But before we get into that meat, let's imagine you were building a utopian society in the best of the science fiction tradition. What would it look like?

### **First we learn to fly our cars**

This is the largest segment of the book. Quite natural, considering the name. Josh details out all the ways in which the world was ready to mass manufacture and adopt flying cars from the late 50s onwards. There were designs, there were prototypes and concepts that worked, there was financing, and there were entrepreneurs aplenty.

And yet the flying cars got sidelined.

Is it because of air traffic control being difficult? No, not really. Because the skies are vast and there's plenty of room up there, what with three dimensions to play with and no "road" restrictions.

Is it because of skill to operate? No, not really. Including his personal example of learning to fly, Josh shows how while it is different to driving, the difficulty is not so high either. And while maybe not every 70 year old grandma might fly, they don't need to. The point here is that the marginal effort is still way skewed in the direction of the flying cars.

Is it the cost? Not really. Before the regulatory haul made a meal of plane costs, Josh shows how a Cessna used to cost c.\$30k, not a crazy amount, while today it's closer to 5-10x that. So clearly there's room to move.

Is it safety? No again. "The leading cause of death amongst active pilots is ... motorcycle accidents." I don't know the source of the statistic, but it's true that worst case scenario, aviation seems about as deadly as riding motorcycles.

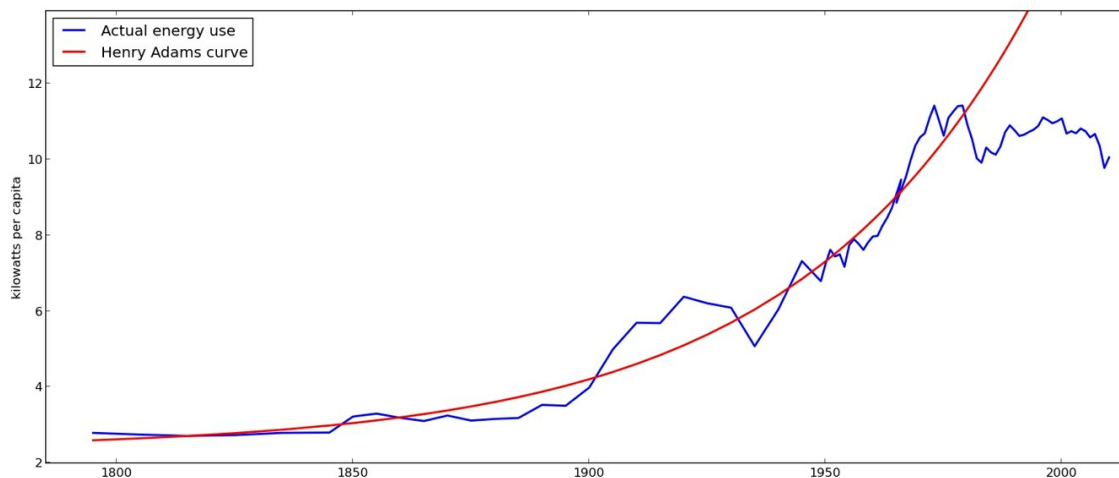
Is it legal liability? Yes to some extent! Anything to do with machines in the air have gotten ungodly scrutiny from the FAA and pretty much forced to shut down, just like what happened with pretty much the entire small aircraft industry.

Part of his argument about the benefits of flying cars is that our effective range can increase from tens of miles a day to hundreds of miles a day. It's almost the same argument made about communication devices and even the internet.

### Then we learn about limitless power

The conventional wisdom goes that it's expensive, and that it's dangerous. Neither of these are inherent properties. Re the danger, Josh shows us how we're overregulating to take away even the slightest chance of something going wrong, and in such pre-emptive regulating, we end up making the technology stuck in stasis.

The point here is that there's been a clear change in our previously-exponential energy usage, what he calls the "Henry Adams Curve" which is a rather clear divergence. The wage line here, you'll note, is the obligatory "wage stagnation" one that I [examined here](#).



*Energy consumption per capita in the US. One kilowatt, of course, equals 8,766 kilowatt-hours per year.*

The same thought process also goes regarding its expense. Part of it is of course the increased regulatory burden imposing hefty costs and, in this case, not providing a single new license for decades! He shows how the construction costs of US nuclear power plants went up by 7x between 1980 to 1995.

Even in the Sixties we knew how to make artificial radioactive isotopes of various elements by exposing them to neutrons. Josh discusses this, and also how beta

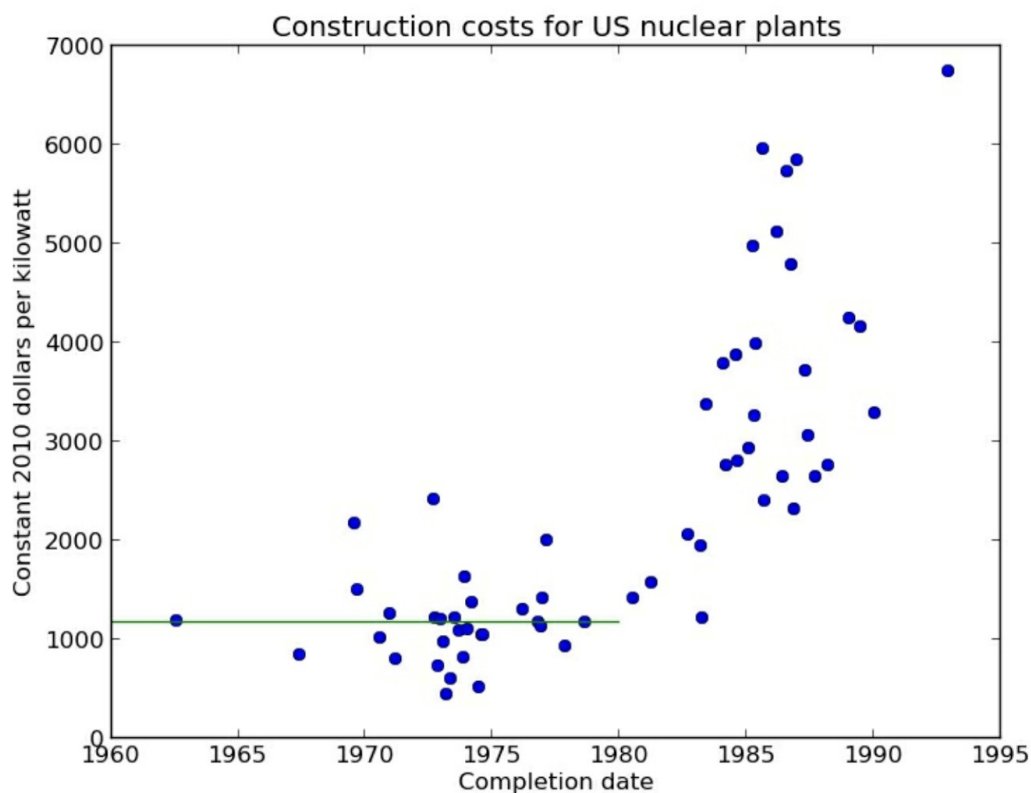


radiation, produced this way, can be blocked by a couple inches of water or a thin sheet of metal.

He mentions this as an example of how there's a colossal Failure of Nerve with respect to trying to find a Moore's Law equivalent for energy. The fact that nuclear fuel is more energy dense and can product high energy neutrons isn't enough of a reason to not follow through on at last examining it.

Josh also examines whether we're just not going after nuclear power because our knowledge is far ahead of our ability to control it, and comes to the conclusion no. He suggests

Nuclear Physics lacks a coherent theoretical foundation that would permit us to analyze and interpret all phenomena in a fundamental way; atomic physics has such a formulation in quantum electrodynamics, which permits calculations of some observable quantities to more than six significant figures. ... Two of the leading theories, the liquid drop model and the nuclear shell model, each have areas where they give good predictions in agreement with experiment.



*Construction costs for nuclear power plants skyrocketed after the establishment of the DOE. Horizontal line is the pre-1980 average, at \$1175/kW. At today's prime rate of 3.25%, that would amortize to about half a cent per kWh. Note that regulation also considerably bloated construction times.*

The problem with nuclear energy tech is not just that it's expensive and has seen pretty insane regulatory headwinds. For example, France doesn't share the worries that US has, and generates 70% of their energy from nuclear.

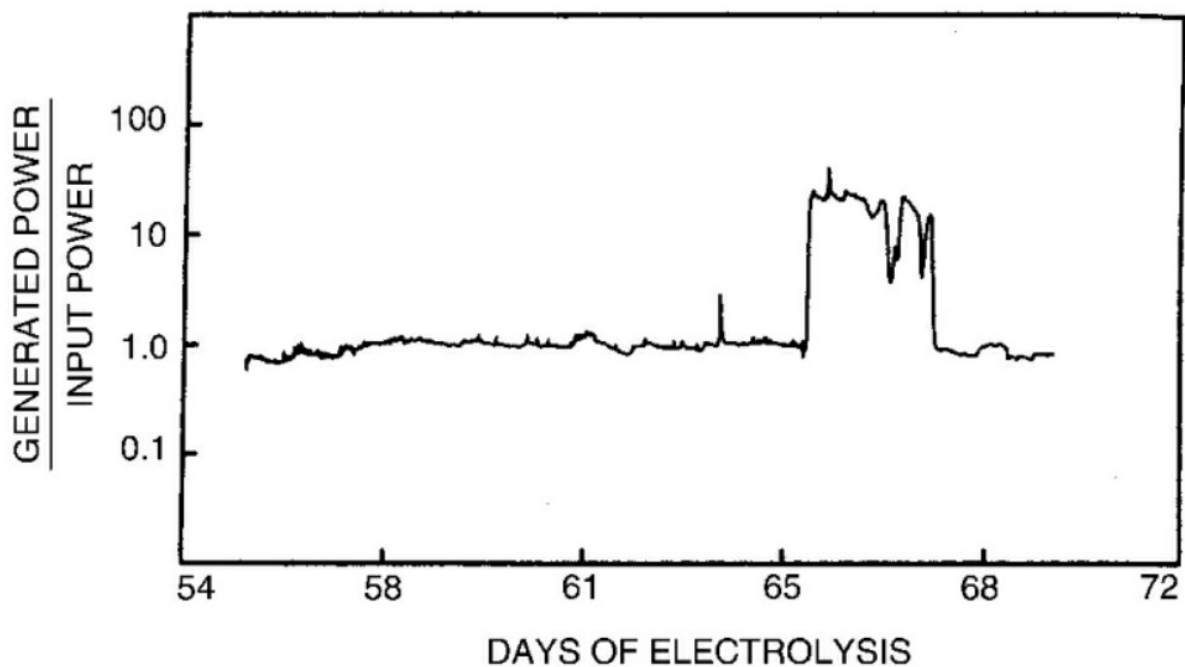
It's that with US not leading the charge, and other countries alongside it, there has been minimal ability to come down the cost curve as drastically as we'd like. It also means that there's not been enough efforts to actually try new techniques in nuclear power generation.

The efforts have shifted from engineering decisions that need to be continually improved, to regulatory decisions that require extremely stringent controls.

It is important to say here though that this seems to be (slowly) changing now. There are startups getting funded which try new forms of reactors, including fusion, like TAE Technologies, Tokamak Energy, Commonwealth Fusion or General Fusion. Is this the beginning of a new expansion? We don't know, but it's true that it's taken a few decades of stasis to get here!

### **And moving on to the beautiful promise of Cold Fusion**

The story here starts with Fleischmann and Pons who invested \$100k of their own money to experiment on cold fusion in 1985. Needless to say, since we don't have cold fusion yet, they ran out of money and didn't have much to show for it, except for this one, anomalous, result.



They were not the first, and they weren't the last. The Department of Energy dug into the phenomenon and found nothing. Without going into detail about all the other trials of the technology, suffice it to say cold fusion quickly became synonymous with a massive embarrassment and pretty much the equivalent of magical pixie dust.

One major problem with the state of cold fusion research today is that while there is a small cadre of smart and careful scientists slowly making small advances, there is also a substantial number of flakes, crackpots, and mountebanks, making outrageous claims—and regularly being discovered to be self-deluded or frauds when serious testing is done on their apparatus. Furthermore, due to the high difficulty of replication, it is all too easy for an honest but inexperienced researcher to fool himself into thinking he has positive results when he doesn't. It wouldn't be at all surprising if the cold fusion literature were at least as unreliable as, say, the biotech literature, where studies have shown that only about 20% of published results could be replicated.

But Josh cites Arthur C Clarke, Julian Schwinger, Brian Josephson (latter two being Nobel winners) and several other notable physicists as supporters of research into cold fusion, lamenting the fact that we don't spend nearly enough time or attention on the topic.

But by 2008, there had been over 300 published replications and verifications of cold fusion phenomena by scientists around the world.

Josh makes the claim that cold fusion has been unfairly maligned, that there have been successful small scale tests, and that it needs more resources put against it. Considering the evidence, albeit circumstantial, it's not entirely unreasonable!

### **And we build palaces, atom by atom**

Richard Feynman talked about nanotech in a particular fashion. He thought the key to successfully creating nanotech was to create machines that could make machines a size order below, and continue pushing the boundaries of what's possible to build in that direction. While it's not exactly the same as a von Neumann replicator, there are similarities. And Josh wholeheartedly agrees here.

Josh spends a lot of pages on the technological barriers in place that stops us from making von Neumann nanorobots. The dream is to be able to make entire machines in nanoscale, swarms of small robots that are able to help solve everything from health issues inside your body to building magnificent machines and new materials like we've never seen before.

For instance, Josh details the problems that remain in place - reaching the atomic scale means that we have to treat materials as discrete atomic particles rather than a continuous metal, forces impacting on particles changing from gravity to adhesion, increased heat dissipation, and in case that's not enough, quantum mechanical tunnelling.

But still, the blame that Josh lays here is on the centralised research bureaucracy that wrote off these ideas. He coins the term Machiavelli Effect as a term for when older, established folks don't want to effect change while the beneficiaries of said change don't push hard enough to make it happen.

It's worth saying the Josh is deeply involved in the field of nanotechnology. He founded the sci.nanotech Usenet and moderated it for 10 years, and was the founding chief scientists at Nanorex Inc. He's also developed several ideas that sound like they should

be in a Crichton novel but are actually real, like the utility fog. So it's one of those cases of listening to the expert in the field express optimism that a super difficult problem set will be cracked if we just apply ourselves.

## **Building the Tower of Babel and 100km space piers**

One of the best detailed part of the book brings a few of the technologies together and suggests how we can build cities that are an entire single building. Josh goes through the calculations, how we can house 40 million people in a tower ten miles high with a footprint of a square mile.

A ten-mile tower might have a footprint of a square mile and could house 40 million people. Eight such buildings would house the entire current population of the United States, leaving 2,954,833 square miles of land available for organic lavender farms.

The benefit of a flying car is that the third dimension is our friend. The benefit of nanotechnology is that building a ten mile high tower is within our grasp. The benefit of nuclear energy is that we won't have to worry about the energy requirements to build any of that.

That's the lesson of the book. The type of reality we *could* build if we were to really work and solve the problems in front of us. If it helps, I read this while listening to Stairway to Heaven, which helped with the mood.

In exactly the same way, were we to solve all these problems, we could create an aircraft that has a ten-mile wingspan. Perhaps shaped like a manta ray. With chords give miles at the center, a mile thick, with enough volume to house 10 million people at 12k square foot each. It would have 250 levels of roadways, 50k elevators and more. A flying city.

We could even build space piers that stand atop 100km towers, where we could launch items into orbit and beyond with a railgun. If we're able to manufacture diamonds aplenty as a building material, or maybe something even better, this is something we could accomplish.

Flawless diamond, with a compressive strength of 50 GPa, does not even need a taper at all for a 100 km tower; a 100-km column of diamond weights 3.5 billion newtons per square meter but can support 50 billion. Even commercially available polycrystalline synthetic diamond with advertised strengths of 5 GPa would work.

The words "even commercially available" is what stands out to me above.

## **And a consequent dilution of culture**

Green fundamentalism gets its fair share of slapping around in the book. Josh shows again and again how they overreact to the potential environmental hazards of nuclear energy and flying cars. This also drives ergophobia, which in the tome means the fear of increasing our energy usage. The level of attention the Greens get in the book seem overplayed considering their impact has been, if anything, rather muted on a macro scale.

Josh also talks about our cultural zeitgeist is one of winner-take-all decisioning, where we're convinced that the way to get ahead is to increase our share of the pie, rather than increase the pie itself. Here he draws on Peter Turchin's idea that when an empire gets large enough it stops worrying about existential threats from outside and rather focuses on internal competition. The thesis being that, just like Ozymandias in the Watchmen, we need an external threat to make us all point in the same direction!

Technologies that provoke antipathy and promote discord, such as social networks, are the order of the day; technologies that empower everyone but require a background of mutual trust and cooperation, such as flying cars, are considered amusing anachronisms.

That's a highly summarised flavour of the book. There are deep and technical discussions about wingspans and lift with the cars, energy potential in nuclear power, dollar/kg cost of launching payloads to orbit and its cost curve, newtons/sqm calculations of a diamond column's weight support, half lives of various nuclear isotopes, and many many more. The organization of the book presumably made a whole deal of sense to the author, but I had to flip back and forth a ton on my kindle to get here.

## II

It's a compelling story. We have clear narratives in place that tell us that we're in a place of technological stagnation, political polarisation, economic growth bifurcation and general malaise affecting a large swath of the economy. But the question I kept asking is why these technological developments seem so highly contingent in the first place. It's difficult to blame regulations and bureaucracy as the sole reason why these technologies haven't come about. As I'd written about in [my examination on wage stagnation](#), while regulations did rise in several sectors, including at the federal level, the time also saw fat profits coming in the private sector who were subject to said regulations.

If we are to believe that the Great Stagnation was a result of us getting in our own way, we have to address the two major issues that Josh identifies – 1) the regulatory burden that we impose on anything resembling new innovation, and 2) the lack of ambition/ nerve/ interest in continuing our search for the next avenues of growth.

### **Onerous regulations**

The growth in regulations was not a binary event. It happened gradually, in stages, and so we should expect to see the chilling effects it had also come about gradually. The regulations didn't happen in a vacuum.

For instance, the regulations to do with flying vehicles is examined in great depth, especially poetically because the author himself is a pilot. It shows, for example, how there is a 175,000 page long regulatory manuscript from the FAR which even has rules specifying who can do paperwork about the maintenance of an aircraft.

But even here, it didn't go from 0 to 175,000 in one shot. There's a historical contingency in this growth. Even allowing that this is insane, and it so clearly is, what's unclear is why, like every other industry on the planet, this also didn't see lobbying to

make life easier. After all the Dodd-Frank bill was incredibly complex too. It created several agencies and added a multitude of requirements on the existing financial institutions.

But it still hasn't stopped companies from forming or technologies from developing. Unlike what we see in flying cars, which stayed in limbo for a few decades, people keep throwing money at folks who think they can solve the problem.

So this is not what we see generally. The industries that Josh argues should have been pursued pretty much look like they hit a brick wall. It didn't go from "approval takes a month" to "approval takes six months", but "approval takes a month" to "let's just not do this at all". That's a rather drastic change.

None of this is to say regulations don't matter. They absolutely do and they slowly strangle new industries through sheer intransigence. The rise of fintech in London came about through deregulation, as a case in point of how it can very well be stifling in its effect.

What I'm arguing is that when there are clearly laid out paths by multiple interested parties, you need more than creeping regulatory rise as a cause for their death. You need to explain why no well meaning philanthropist or venture capitalist didn't burn a few hundred million in chasing the dream.

You would also have to explain why companies like Lilium seem to be coming about now, and the whole VTOL movement, and companies making supersonic crafts like Boom, since regulations are *still crazy*.

The only way that argument makes sense is if we insist that the industries were all at the cusp of unprofitability, and the regulatory shifts towards becoming just that much harder just pushed them over the edge. And believing that multiple technologies were all simultaneously teetering at the edge of commercial viability seems a tad inappropriate.

Sure you can make that case maybe for nuclear power, since around the same time we started getting interested in solar. But what about healthcare? That's insanely regulated and continues to be even more rigidly regulated. Like a twitchy mob boss if you so much as look at it sideways it gets all uppity and increases its prices for the nearest drug or procedure by 300%.

It's also insanely bureaucratised with enough paperwork to make even the Pale King happy. But there is still investment going into the sector. There's still innovations being uncovered and there's still progress, even if parts of it are slower than what we would like.

It means there's something else going on behind the scenes. Somehow the institutions that we set up end up becoming nooses around our neck as opposed to helping us grow. It's a bigger problem to do with increasing [organisational inertia](#) rather than an easily solved choke point.

There has to be a reason why the operators, the financiers, the venture capitalists, the large tech conglomerates, none of them saw fit to pursue the dreams that are laid out in this book.

## Failure of nerve and imagination

And here's the second, and arguably more interesting aspect of the book. Josh coins the phrase "Failures of Nerve and Imagination" which he specifically speaks about the problem that those who oversee science and research funding have.

One of the great tragedies of the latter Twentieth Century, and clearly one of the causes of the Great Stagnation, was the increasing centralization and bureaucratization of science and research funding. This meant that Failures of Nerve and Imagination, which are particularly strong among bureaucrats, instead of merely causing incorrect predictions from pundits, caused resource starvation and active suppression, and became self-fulfilling prophecies.

But isn't this typical of bureaucrats and overseeing bodies always? How recent a development is this that the Failures of Nerve and Imagination seems to be temporally aligned in a particular time period, or sectorally in particular sectors? That seems like a bit of a stretch.

Also, this is a far sweeping assertion about lack of funding. But it's just not true. We spend billions on insane research. Bear in mind that the CIA wasted money on equipping bats with bombs, mind control and LSD. Are you really claiming that nanotech was where they drew the line?

Of the billions appropriated by the military to research the most absurd ideas on the planet, they didn't want to actually explore the slightly-less-absurd end of the spectrum?

It especially falls apart because one of Josh's key ideas is the Machiavelli Effect. It's essentially the idea that elites are resistant to change when there's a chance it could hurt the elites. So for example he says:

Centralized funding of an intellectual elite makes it easier for cadres, cliques, and the politically skilled to gain control of a field, and they by their nature are resistant to new, outside, non-Ptolemaic ideas. The ivory tower has a moat full of crocodiles.

But if that's the case, then shouldn't the more opaque and byzantine institutions either do only the most run-of-the-mill research? If they can put radar and explosives on bats, why can't they dig deeper into nanotech?

Another one of the issues here is on funding from private sources. Here the Machiavelli Effect that Josh coins is applied to folks like VCs. But again, VCs would never fund something like this anyway. VCs fund research where it can return tangible results in a predictable fashion across portfolios. For instance, they might fund 10 different drug molecules because they know one of the ten will break through. There's just no equivalent in nanotech of the sort Josh talks about.

## III

My conclusion is that while the effects that Josh describes are undeniably true, and the causes he describes play only a supporting role, there has to be a more systemic explanation.

For instance, we should know why despite [nuclear energy being a mainstay in France](#), that liberal bastion, that provides 70% of their needs, it still feels like it's in a technological and political stranglehold in several places in the world. The same in [South Korea too](#), where they saw highly successful nuclear power development for a while before plateauing.

And if we had pockets where nuclear power was used successfully, such as in the Navy, why didn't it expand to more civilian areas? We managed to get the GPS to become ubiquitous, but a cheap and clean energy source was somehow taboo?

Bear in mind that China, as the latest superpower, is also [building plenty of power plants](#), with almost 5% of the total electricity in China coming from nuclear energy as of 2019. Once again, there does seem to be movement in this domain, just not within the United States. We should know why that is.

We should also know why despite the military spending billions on building robots that can run and even self driving cars, they didn't fund nanotech to a degree that would make Josh happy.

We should know why flying cars became scary around the 60s and 70s as opposed to other things, like miniaturising supercomputers, or Thalidomide, or even rocket science.

So if the question is why aren't we doing anything to advance the state of these affairs, that's begging the question. Since my answer is that we actually are. Badly, grasping at fake straws at times, but we are. Our [scientific productivity has been falling off a cliff](#) for many disciplines, but we also [dedicated two decades](#) to the study of a [biological process](#), under duress, which resulted in our ability to fight the pandemic in 2020. [Isolated narratives of progress](#) makes it sound like what innovations we get are the rare oasis amidst a desert, but they're not. Our ability to push forward multiple lines of inquiry are what lets us solve some of them, and for multiple dominoes to seemingly fall at the same time.

I also looked around to try and figure out if there had been any research done to try and identify how and if regulatory underreach leads to similar counterintuitive outcomes too.

There was an empirical paper I found answering this question, which looked at the outcomes of regulation called "How deregulation can become overregulation", which had the following line:

Functional regulation is a deregulatory measure, giving the organizations the responsibility to implement systems that result in safe operations, while governments only are to verify that the organizations have implemented systems. Contrary to the deregulatory intentions, the organizations implement rules, clutter and red tape that often not contribute to safety, and even continue to grow. Many organizations have tried to simplify their safety management systems, but still have ended up with at least as many procedures as before (Power, 1999).

...

Paradoxically then, deregulation seems to lead to overregulation.

They describe the condition as overregulation, since the internal regulation is detailed and overachieving on the limit to contradict its objectives.



Though it's an empirical study of Norwegian coastal cargo and fish farming, the implications definitely don't seem limited to that instance.

So if regulation by itself can't help explain the nosedive that our innovative spirit took, can something else? Something that created our failures of nerve and imagination? It's unclear that there were enough external events that made this happen either. For instance, 1970s saw an oil crisis. Shouldn't this have acted as a catalyst for us to examine other energy sources? Especially if it were one of the key events that led us to re-examine our entire very energy dependence and flattened the Henry Adams Curve.

If the argument is that energy becoming scarce combined with Green fundamentalism to push us into conservation mode, and this happened a few decades before Climate Change even became a household worry, then you have to also ask why this didn't lead us into a renaissance of looking at cleaner sources.

My read is that there's a tangled web of motivations here. It's like a simple graph with seventy variables, all of which have relationships with each other. There are plenty of stories that can fit the observed pattern, and plenty of causal stories that can be drawn. To be so convinced of a conclusion here seems, to me, a tad hasty.

Every large company and organisation goes through the same calculus, is the effort worth the payoff? And that answer comes from assessing the costs of the regulation alongside the rest, and analysing whether it makes sense to follow through. None of them usually just focus on the left hand side, the regulatory side, and walk away.

In industry, for the longest while, technological advance seemed completely tied to manufacturing capacity. While the creation of the car is an undeniable technological advance, it became a household tool only when Henry Ford discovered the assembly line process and 4x-ed the productivity. The reflexivity in between technology per se and its industrial applications also seems applicable to nanotechnology and space exploration that Josh writes about. There's a long road between being a curio in a CIA lab and something that exists in the wider world, and that chasm isn't bridged easily just through willpower. It requires an entire substrate to already exist.

Also, energy demand is elastic, as discovered first by Willian Stanley Jevons in 1856 when he wrote *The Coal Question*. He noted, as written in the book:

...when the steam engine became more efficient with the introduction of James Watt's separate condenser (and many other improvements), the amount of coal used in England grew rather than declining. What this meant was, of course, that people were using the new, efficient Watt engines to do many more things than they had been using the older, less efficient, Newcomen engines to do.

As we find more efficient ways of doing things, we start doing many more things. The best example here is of course electronics, which has become ubiquitous just as it has declined in cost.

It's not a counterpoint per se, but the mood affiliation gets a bit clearer. If you're proposing that our utopian future Jetsons life was derailed by those annoying Greens, government bureaucrats and silly regulators, some of that ire would also spray onto unsuspecting targets like the "don't eat too much fat" lobby.

The impression I get from reading the book is that Josh is quite upset that we're not more engaged and optimistic and confrontational about the opportunities that exist in front of us. Which is fair. What I think he misses is the fact that our efforts to grow is often contingent upon what has already been done, what is available, what society expects from us, and what those [other geniuses around us](#) do. While there are legitimate gripes about our ability to get anything done, I'm not sure throwing the technological growth baby with the stagnation bathwater actually explains much. The explanations Josh advances all seem accurate, though peripheral, to the issue that it *seems like* our progress in technology has stagnated. To answer it we have to a) prove that technological stagnation is real, and not just in isolated pockets like flying cars or nanotech, and b) identify a cause for it that goes beyond our search for energy efficiency.

As humans we like all encompassing explanations for large macroscale phenomena. While this stands us in good stead when it comes to physics, it has led us astray when it comes to economics, politics and sociology. So while I reject the overarching explanation for us being a bit too chicken to do anything, I do think that we could do more. Should we push our advances in more areas? Absolutely.

And we're doing that. To quote myself:

Equally interesting is that the UK is spending \$250m to help bring about the world's first fusion power plant. Again, the success here is not the point as much as the effort. This follows the footsteps of China and a European one. And there are also several companies around the world that are trying to make nuclear fusion reactions more portable. There's CFS that's MIT backed, TAE in California, First Light Fusion in Oxford, Tokamak Energy in Oxford, and the world's largest fusion project in ITER in Southern France. This is a technology that's always been a couple decades away and still might be so. But anything aiming to bend the Adams curve is worth noting.

So the question is less one of if we will ever fulfil our potential, and rather one of what all do we need to do to ensure that we *can* fulfil our aspirations. And that's a much harder question.

## Coda

Two other tidbits that I liked in the book that I want to mention. 1) Nuclear physics appropriated the term nucleus from biology where it was used for the central organelle in a cell since 1831. Same regarding the word fission. And 2) Almost in passing Josh also entertains the idea that "fat is unhealthy" diet is incorrect, and shows plenty of insight from his personal experiments. However the study he draws upon, from Gary Taubes, has rather significant flaws that Scott Alexander has [pointed out](#).

The last point is one that was pointed out by Patrick OShaughnessy on [Twitter](#).

We wanted flying cars and all we got was satellite internet blanketing the world, supersonic jets, 2-day vaccine development, a budding genomics revolution, a \$1T self organizing internet native currency, and AI advancing at insane rate, doubling compute demand every few months.

In many ways this was a book that explored several themes that have also come up in my search – the search for a cause of the [Great Polarisation](#), the search for [why wages stagnated](#), the search for why [progress seems unbounded](#) on the one side and extremely bounded on the other, the question of why [organisations seem to suck so often](#), and the question of [why hierarchies are amazing](#) to get things done but frustrating to get any one thing done.

Our belief in growth comes from an inductive process. A belief that as it has come before, where humanity solved its largest struggles, so we will do again. The pessimists similarly demonstrate an [incredulity](#) on whether we will ever be able to achieve such heights as we have before. Some look at the institutional setup that we have built up and see masses of bureaucracy and the mediocratisation of our entire species. Some look at that setup as the inevitable consequence of success. After all, we can't very well organise 7 billion people to do anything without some level of bureaucracy.

I think the takeaway from the book, to me, is that there are plenty of roads not taken, and the avenues for exploration remain as fertile as it ever was. We should be optimistic, not pessimistic, after reading it. It's one of the books that's gotten me to learn and think about a wider variety of subjects in more depth than I would've guessed. Sometimes when it looks like we're just treading water it's because we're building platforms that are needed for [the next S curve](#). I think everyone should get a hold of the book and read it, not as a narrative of pessimism, and not as a barometer of how far askew we've come from some original goal, but rather to see how much remains to be built, and to get inspired.

# 'Without Gloves', 'Without Grease', 'The Great Game of Politics' and 'Political Behavior' by Frank R Kent

Why read Frank Kent? He worked with Mencken on the Baltimore Sun, wrote in Mencken's 'American Mercury', did the standard book on practical politics of the 1920's-30's, 'The Great Game of Politics', had a column in Time magazine till he died in the 1950's. His best was written in the twenties and thirties and he never claimed to write for all time.

Politicians paraphrased him. When FDR spoke against 'Fat Cats', I thought he was talking about rich people generally. In Frank Kent 'Fat Cat' is a term of art for

'Men of large means, who having reached middle age, having achieved success in finance or business and there being no further sense in the mere piling up of millions, develop a yearning for some sort of public honor or prestige.

...Such men are known in political circles as 'Fat Cats and they are as welcome in the organization as flowers in May. This can be accepted as fact- any 'Fat Cat' able and willing to spend as much as necessary can get what he want in state politics, provided he has not exercised the excessively bad judgment of picking a party hopelessly in the minority in his state. In other words, while under normal circumstances the nominations for higher offices go to the men who have come up the escalator, when a 'Fat Cat' appears on the horizon, there is a rush to take him'.

So FDR was partly attacking the rich as such and partly attacking rich men who wanted to get into politics, say Orange Hitler 2015.

When Churchill said 'Democracy is the worst form of government except all the others', he was paraphrasing Frank Kent's 'Machine Politics is the worst form of government except all the others'.

In 1923 when Kent wrote 'The Great Game of Politics', there were States Rights D from the Southland, big city D like Tammany, Progressive R like Teddy Roosevelt, and machine city R. Reservation Indians, Southern blacks, new immigrants and felons had no vote.

Before reading Kent, I thought FDR was a big government D from the start. In Kent, FRD won 1932 as a small government States Rights Democrat against the spendy

busybody big government Hoover. FDR started out firing 100K R federal employees, then switched gears, hired 500K D federal employees and got millions of voters taking federal money for declaring loyalty to D FDR. This was the death of the old R machine, which was heavily dependent on Federal patronage- they had to turn into States Rights supporters, their enemies since the Civil War. Also, since Hoover disliked black people and told them so, the black vote had kind of moved to D in 1932. Even before Hoover, R had been losing black support, mostly because of R white politicoes in the South.

'The thing that keeps him [the black voter] out, as much as anything else, is the attitude of the white R politicians in those [Southern] states. These men all hold Federal offices. They are not in politics for their health. They are in it solely and entirely for the federal patronage. It is a business for them. Distinctly it is not their interest to build up a virile, fighting R party in these [southern] states. Such a party would develop competition for control and candidates for their jobs. That is the last thing that they want.'

As it became clear that FDR actually wanted black voters enough to support black people getting federal jobs, the black vote moved heavily D, though I think it was still competitive until Affirmative Action.

But these are sidelines for Kent. His main point is that machine politics is the fault of voters not voting in primaries, and that bad as machine politics is it's still better than anything else in the real world so far. Again, he expects high levels of humbug in all political speech, -bad, but better than no free speech:

'Primaries are really the key to politics. There is no way for party candidates to get on the general ballot except through the primaries. Control of that gate in any community means control of the political situation in that community... The potent thing about these machine men is that they vote .. But the overwhelmingly big thing is that they are primary-election voters, not just general-election voters. . . .It is an undeniable fact that, all over the country, the great bulk of men and women who do vote are practically dragged to the polls by the machine workers, and it is a good thing for the individual communities and the country as a whole that this is true. If the machines did not work and pull and haul to get enough voters out in the primaries to put over the machine candidates, the present state of indifference and ignorance of the average citizen would permit out candidates to be chosen for us by the freaks and fanatics who abound in every community, and are constantly and zealously stirred to political activity in behalf of their half-baked schemes for saving the world.'

'Humbuggery in Every Campaign' and 'Why the Newspapers do not print all the facts' are chapter headings in Kent.

'Men do not blurt out everything in their souls when dealing with other men. It would be a terrible world if they did. We are all humbugs to a certain extent. The difference is that when a man -even the best and highest types we have- enters public life and begins to seek votes, he is forced into a position where he must humbug a great deal more than in private life. He is faced with the necessity of making a favorable impression upon a large number of widely scattered, highly diversified groups of voters, notoriously swayed by prejudice, who are being fished for from every conceivable angle by his opponent. He has to watch his step with the utmost care. He has to hide facts, in themselves harmless, and he has to guard and weigh his words on the minor issues and on matters of no real importance, for fear a blast from the opposing camp that will scatter his carefully herded voters like a covey of partridges at the sound of a gun. Moreover, he is forced to fit his public views and utterances into the prejudices of the elements back of him and, also, to some extent at least, into the notions of the men who have put up the money for his campaign.'

Kent wrote when D and R were both about even. Things changed when FDR got a huge chunk of voters on federal patronage, and since Lyndon Johnson doubled down R has been a shadow of a party. If a controlled press and fifty million ringers give us a one party state things will change again.

# Working With Contracts: What Law School Doesn't Teach You by Charles M. Fox

Contracts is one of those areas that I always figured I ought to study, at least enough to pick up the basics, but never seemed either interesting or important enough to reach the front of my queue. On top of that, there's a lot of different angles from which to approach the subject: the law-school-style Contracts 101 class covers the legal principles governing contracts, the economists' version abstracts away the practical specifics and talks about contracts in game-theoretic terms, more business-oriented books often focus on negotiation, etc.

["Working With Contracts: What Law School Doesn't Teach You"](#) is about the practical skills needed for working with contracts on an everyday basis - specifically the sort of skills usually picked up on the job by young lawyers. It talks about things like what to look for when reviewing a contract, how to organize contracts, why lawyers use weird words like "heretofore", various gotchas to watch out for, etc. It assumes minimal background knowledge, but also includes lots of technical nuts and bolts. In short, it's the perfect book for someone who wants a technical understanding of real-world contract practice.

This post will review interesting things I learned from the book.

## Background Knowledge

First, some very brief background info, which the book itself mostly assumes.

Legally, in order to count as a "contract", we need four main pieces:

- Offer: someone offers a deal
- Acceptance: someone else accepts it
- Consideration: both parties gain something from the deal; it's not a gift
- Mutual understanding: both parties agree on what the deal is and the fact that they've agreed to it

A Contracts 101 class has all sorts of details and gotchas related to these. Notice that "signature on a piece of paper" is not on that list; e.g. oral contracts are entirely enforceable, it's just harder to prove their existence in court. Even implicit contracts are enforceable - e.g. when you order food from a restaurant, you implicitly agree to pay for it, and that's a legally-enforceable contract. That said, we'll focus here on explicit written contracts.

Once formed, a contract acts as custom, private law between the parties. Enforcement of this law goes through civil courts - i.e. if someone breaches the contract, then the counterparty can sue them for damages. Note the “for damages” in that sentence; if a counterparty breaches a contract in a way that doesn’t harm you (relative to not breaching), then you probably won’t be able to sue them. (Potentially interesting exercise for any lawyers in the audience: figure out a realistic contractual equivalent of [Newcomb’s problem](#), where someone agrees to one-box on behalf of someone else but then two-boxes, and claims in court that their decision to two-box benefited the counterparty rather than harming them. I’d bet there’s case law on something equivalent to this.)

Note that this is all specific to American law, as is the book. In particular, other countries tend to more often require specific wording, ceremonial actions, and the like in order to make a contract (or component of a contract) enforceable.

## What Do Contracts Do?

The “functional” components of a contract can be organized into two main categories: representations and covenants. A representation says that something *has happened* or *is true*; a covenant says that something *will happen* or *will be true*.

Some example representations:

- ABC Corp signs a statement that they have no pending lawsuits against them.
- Bob signs a statement that the house he’s selling contains no lead-based paint or asbestos insulation.
- Carol signs a statement that the forms she provided for a mortgage application are accurate and complete.
- Title Corp signs a statement that there are no outstanding mortgages on a piece of property.

Nominally, each of these is a promise that something is true. However, that’s not quite how they work *functionally*. Functionally, if a counterparty acts based on the assumption that the statement is true and is harmed as a result, then they can sue for damages. In other words, when providing a representation, we provide **insurance** against any damages which result from the representation being false. Bob may not even have checked that the house he’s selling contains no asbestos, and that’s fine - *if* he’s willing to insure the counterparty against any asbestos-related risk.

This idea of insurance becomes important in contract negotiations - there’s a big difference between e.g. “no environmental problems” and “no environmental problems *to the best of their knowledge*”. The former insures against any environmental problems,



while the latter insures against any environmental problems which the signer knew about at time of signing. One puts the duty/risk of finding/fixing unknown problems on the signer, while the other puts it on the counterparty.

The other key thing to notice about representations is that they're *as of the signing date*. When Bob states that his house contains no asbestos, that does not insure against the house previously containing asbestos or containing asbestos in the future. It only needs to be true as of that one moment in time. This becomes relevant in complex multi-stage contracts, where there's an initial agreement subject to a bunch of conditions and reviews, and the final closing comes later after all that review is done. For instance, in a mortgage there's an initial agreement subject to the borrower providing lots of forms (credit check, proof of income, proof of insurance, etc...), and the final contract is closed after all that is reviewed. In these situations, the borrower usually makes some representations early on, and then has to "bring down" the representations at closing - i.e. assert that they're still true.

While representations deal with past and present, covenants deal with the future. They're the classic idea of contract provisions: precommitments to do something. Some examples:

- ABC Corp agrees to not sell the machinery they're leasing.
- Bob agrees to not use any lead-based paint on the house he's buying.
- Carol agrees to maintain minimum levels of insurance on the house she's mortgaging.
- Monitoring Corp agrees to alert Bank if there is any change in the credit rating of Company.

These work basically like you'd expect.

Representations and covenants often run in parallel: a representation that X is true will have a corresponding covenant to make X continue to be true in the future. For instance:

- ABC corp states that they do not currently have any liens on their main plant, and agrees to not create any (i.e. they won't borrow any money with the plant as collateral).
- Carol states that she currently has some level of insurance coverage on her house, and agrees to maintain that level of coverage.

This is mainly for contracts which will be performed over a long time, especially debt contracts. One-off contracts (like a purchase/sale) tend to have relatively few covenants; most of their substance is in the representations.

## Parallels to Software Development

Representations and covenants seem pretty straightforward, at least conceptually. One is insurance against some fact being false, the other is a precommitment.

The technical complexity of contracts comes from the interplay between two elements.  
First:

“The goal of a contract is to describe *with precision* the substance of the meeting of two minds, in language that will be interpreted by each subsequent reader *in exactly the same way*.”

In other words, we want no ambiguity, since any ambiguity could later be used by one of the parties to “cheat” their way out of the contract. This creates a headache very familiar to software developers: like programs, contracts mean exactly what they say. There is no “do what I mean” button; we can’t write something ambiguous and rely on the system to figure out what we meant.

Second: we don’t have perfect knowledge of the future. When making a precommitment in a contract, that precommitment is going to operate fairly mechanically in whatever the future environment looks like. Just like a function written in code may encounter a vast space of unusual inputs in the wild, a precommitment in a contract may interact with a vast space of unusual conditions in the wild. And since we don’t know in advance *which* conditions will be encountered, the person writing the code/contract needs to consider the whole possible range. They need to figure out, in advance, what weird corner cases *could* arise.

Put those two pieces together, and the picture should feel *very* familiar to software developers.

The result is that a lawyer’s job ends up involving a lot of the same pieces as a software engineer’s job. A client/manager says “here’s what we want”, the lawyer/programmer says “ummm I don’t think you really want that, because <problem> happens if <circumstance>”, and they go back-and-forth for a while trying to better define what the client/manager really wants. An example from the book pictures a lawyer reviewing a contract with a client (simplified slightly by me):

“Lawyer: This is a covenant that restricts your business from incurring debt...

Client: That’s fine, we don’t plan to use any bank financing.

Lawyer: Well, the definition of “debt” used is very broad. For instance, it includes payment plans on any equipment you buy...

Client: Well, we can add some room for that.

Lawyer: How much room do you need?

Client: Based on our current needs, less than \$1M at any given time.

Lawyer: But if that new plant you were talking about gets off the ground, won't you need to buy a bunch of new equipment for it?

Client: Good point, we'd better ask for \$5M..."

This could go on for a while.

Despite the parallels, lawyers are not very *good* software engineers, in general. The most common solution to the sorts of problems above is to throw a patch on it, via two kinds of exceptions:

- Carveouts: action X is generally forbidden, except for special case Y.
- Baskets: action X is generally forbidden, except in amounts below some limit (e.g. the \$5M limit in the example above)

Over the course of negotiations, patches are layered on top of patches. An example from the book:

"Little Corp may not transfer any Shares during the term of this Agreement, except for (i) transfers at any time to its Affiliates (including, without limitation, Micro Corp) other than Medium Corp, and (ii) so long as an Event of Default attributable to Big Corp shall have occurred and be continuing, transfers to any Person (including, for the avoidance of doubt, Medium Corp)."

This mess is the contractual equivalent of a series of if-statements nested within if-statements. This is, apparently, standard practice for lawyers.

(Another complaint: in a complex contract, it would not be hard to include provisions alongside the table of contents which nullify provisions which appear in the wrong section. Then people reviewing the contract later wouldn't have to read the whole thing in order to make sure they didn't miss anything relevant to their use-case; it would be the contract equivalent of variable scope. My mother's a lawyer in real estate and wills, so I asked her why lawyers don't do this. Her possibly-tongue-in-cheek-answer: might put lawyers out of business. Kidding aside, the bar association engages in some pretty incestuous rent-seeking, but judges have been pushing for decades to make contracts and other legal documents more legible to non-lawyers.)

## The "Do What I Mean" Button

A contract writer's job is much easier than a programmer's job in one key respect: a contract will ultimately be interpreted by humans. That means we can say the equivalent of "look, you know what I mean, just do that", *if* we expect that a court will actually know what we mean.

This gives rise to a bunch of standard tricks for invoking the do-what-I-mean button. We'll talk about three big ones: materiality, reasonableness, and consistency with "ordinary business"/"past practice".

One important thing to keep in mind while reading these: **when you push the do-what-I-mean button, the contract isn't necessarily going to do what you mean, it's going to do what a jury interprets it to mean.** And juries are notoriously unreliable.

## Materiality

Materiality means ignoring small things. For instance, compare:

- "Borrower shall not default in its obligations under any contract", vs
- "Borrower shall not default in its obligations under any material contract"

The first would be breached if e.g. the borrower forgot to update their payment information on their \$10 monthly github subscription, and the payment was late. The second would ignore small things like that.

In general, materiality is relative to the size of the business. A \$100k oversight would be quite material to most small businesses, but immaterial to AT&T. It's also relative to the contract - if that \$100k oversight is directly relevant to a \$300k contract, then it's material, even if the \$300k contract itself is small change to AT&T.

Where's the cutoff line? That's for courts to decide, if and when it matters. That's how pushing the do-what-I-mean button works; you have to rely on the courts to make a sensible decision.

One particularly common usage of materiality: "material adverse change/effect". Rather than saying "X has no pending lawsuits", we say "X has no pending lawsuits whose loss would entail a material adverse effect". Rather than saying "Borrower will notify Lender of any change in their business forecasts", we say "Borrower will notify Lender of any material adverse change in their business forecasts". This way a lender or buyer finds out about problems which actually matter, without being inundated with lots of minor details.

## Reasonableness

Reasonableness is exactly what it sounds like. It's saying something that has some obvious loophole to abuse, then giving a stern look and saying "don't go pulling any bullshit". Example: "Company shall reimburse X for all of X's out-of-pocket expenses arising from..." vs "Company shall reimburse X for all of X's reasonable out-of-pocket expenses arising from..."

Some patterns where reasonableness shows up:

- Reasonable expectations, e.g. "Borrower shall notify Lender of any changes which could reasonably be expected to have a material adverse effect..."

- Consent not to be unreasonably withheld, e.g. “ABC Corp may not X without consent of XYZ Corp, such consent not to be unreasonably withheld.”
- Reasonable efforts, e.g. “Borrower shall obtain X from their insurer.” vs “Borrower shall exert reasonable effort to obtain X from their insurer.”

What would each of these do without the reasonableness clause? In the first case, the borrower could claim that they didn't expect Obvious Bad Thing to impact their business. In the second case, XYZ Corp could withhold consent for some case they obviously don't care about in order to extract further concessions from ABC Corp. In the third case, an insurer could simply refuse to provide X, and the borrower wouldn't be able to do anything about it.

## Behaving Normally

Sometimes a lender or prospective buyer wants to say “what you normally do is fine, so do that and don't go crazy”. Two (similar) standards for this: “in the ordinary course of business” and “consistent with past practice”.

Typical examples:

- “Borrower will not incur any <debt of specific type> except in the ordinary course of business.”
- “ABC Corp will not make any payments to <subsidiary> except in a manner consistent with past practice.”

In general, this is a pretty good way to let business continue as usual without having to go into all the tiny details of what business-as-usual involves, while still ensuring that e.g. a borrowing company doesn't sell all their assets, distribute the funds as a dividend to a parent company, and then declare bankruptcy.

## Remedial Provisions

In general, if a contract is breached, the counterparty can sue for damages. If you want anything else to happen as the result of a breach, then it needs to be included in the contract. In particular, common things triggered by a breach include:

- Termination: counterparty gains the right to terminate the contract
- Acceleration: loaned money must be paid back immediately
- Indemnification: counterparty must be paid for any breach-related damages

The last is somewhat redundant with the court system, but by including it explicitly, the contract can also specify how to calculate damages, how damages are to be paid, caps or exceptions to liability, etc. Rather than leaving such matters to the whims of a court, the contract can specify them.

Termination and acceleration are particularly relevant from a negotiation standpoint - the former for one-shot contracts like sales, and the latter for long-term contracts like debt.

The earlier stages of a complex sale (e.g. a merger/acquisition of a company) involve an agreement to sell *subject to* a long list of conditions being satisfied - i.e. the “due diligence” conditions. If any of those conditions are not met, then the buyer gains the right to terminate the contract - i.e. walk away from the deal. But these things can take months; the last acquisition I saw took around a year. During that time, the buyer may change their mind for reasons entirely unrelated to the seller - e.g. market prices for the seller’s assets may change. The seller wants to prevent the buyer from walking away in a case like that.

This means that the buyer has incentive to ask for very complicated and/or very subjective conditions, to give themselves the opportunity to walk away whenever they want. For instance, if a buyer manages to get a condition which requires “X which is satisfactory *in Buyer’s sole discretion*”, then the buyer effectively gains a blanket option to walk away from the deal; they can always just claim that some inane detail of X is unsatisfactory. (This is a good example where reasonableness can fix the problem.) In particular, if market conditions change, then the buyer may use that option to negotiate more concessions, like a lower purchase price.

Acceleration has a similar effect in debt deals. Nobody ever wants to accelerate debt; it’s a surefire way to end up in bankruptcy court. When a contract breach gives a lender the option to accelerate, what actually happens is that they use that option as leverage to negotiate a new deal. They’ll want a higher interest rate, or a claim on more of the borrower’s assets, or the like.

Takeaway: just because a contract specifies a particular penalty for breach does not mean that the penalty actually happens. Often, the penalty is really used as an option by one party to renegotiate the contract, and provides leverage for such a negotiation.

## Takeaways

Contracts are a lot like computer programs: they’re taken very literally, and they could potentially encounter a wide variety of corner cases in the wild. Together, those two pieces make a contract writer’s job quite similar to a programmer’s job: a client/manager will tell you what they *think* they want, and then you go back-and-forth trying to formulate what they really want.

Compared to (good) software developers, lawyers do not seem to be very good at this; they tend to throw patches on top of patches, creating more corner cases rather than fewer. They don't seem to have even realized that *enforced* scope and modularity are things which one could use in a contract; consequently, every contract must be read in its entirety by anyone relying on it. That puts a sharp limit on the scale of today's contracts.

Unlike programmers, lawyers do have a "do what I mean" button, although its use comes with a cost; it means leaving interpretation to the whims of a court. For many "simple" things, that cost is relatively minor - so contracts can ignore "immaterial" problems, or require "reasonable" behavior, or stipulate consistency with "past practice" and "the course of ordinary business".

Functionally, contracts provide insurance against stated facts being false, and they provide precommitments for the future. They can also stipulate nominal penalties for breach of contract, though in practice these penalties often serve as options to renegotiate (with leverage) rather than actually being used.