# Book Summary: The Righteous Mind

This is my slightly altered version of the Lesswrong post here:

[Book Review: The Righteous Mind - LessWrong](https://www.lesswrong.com/posts/iBDkwir2FsG4DAiDh/book-review-the-righteous-mind-3)

“The Righteous Mind” by Jonathan Haidt is a book about how \*other\* people think. Specifically, how they think about morality and make moral judgments and choices.

The first part of “The Righteous Mind” could hold its own as a book about thinking and rationality.

Like any book about rationality - it needs a central catchy metaphor for how the brain works. Here this metaphor is that of a rider on an elephant. Imagine a man riding on the back of an elephant. The rider can say “turn right!” but if the elephant wants to go left, that’s where they’re going. All the rider can do is say “Yes, we meant to go left. Left is the best choice. Here’s why.” The rider is conscious reasoning, and the elephant is emotions, intuitions and everything else we aren’t even always aware of. Emotions and intuitions make the decision where to turn almost instantaneously, and our conscious reasoning has to go along for the ride.

the elephant (my unconscious emotions) leans towards either “accept” or “reject”. Accordingly, the rider (conscious reasoning) either tries to believe and support the arguments, or fight them internally with everything he’s got. If I want to accept, I ask “can I believe it”; to reject, I ask “must I believe it”. Sometimes, when I’m especially intellectually honest, I can concede an argument and say “this argument is true, but I’m still not swayed about the general issue.” Whew, attack parried.

![Untitled](https://s3-us-west-2.amazonaws.com/secure.notion-static.com/d75cbbc0-c0fa-4dd2-9975-d1ec9460bd65/Untitled.png)

It’s not that whichever view exists is confirmed - this is just a symptom. Rather, whichever view has the intuitive, emotional appeal to someone - that view will be supported and confirmed with reasoning. This might look like confirmation of an existing opinion, but only because this process has run before, the first time the person was exposed to the issue, and resulted in the same outcome - support for the opinion the elephant preferred. So it seems like the existing opinion is confirmed, when instead what is confirmed each time is the elephant’s choice, the emotional valence. It’s motivated reasoning, not confirmation bias.

We’ve all had this experience of trying to convince someone to change their mind, and them just \*not listening to reason\*

. What actually happened is that we were aiming our arguments at the rider, who is not the one making the choice - only the one giving the reasoned justifications. I have a personal rule when I try to convince someone and refute their arguments. If at some point their arguments become \*really\*

bad - I know I’ve lost. Because then I know it’s not really the arguments that determined their opinion. It’s something else that my arguments cannot reach. As the saying goes, “you can’t talk someone out of something they haven’t been talked into.”

Where does that leave reasoned debate? If the arguments aren’t really what convinces people, what other avenues are there? By this view, if we really want to persuade someone, we should definitely NOT say “Look at this table, the benefits of the policy exceed the costs, you should support.” Instead, we should whisper a message to the elephant: “Many people support this policy. The cool kids / popular people / your Team / your Tribe support it. You would gain their favor by supporting it. I’m your friend, you can trust me, and I support you.” Or whatever other signals the elephant picks up on, which are often non-verbal and non-explicit. Advertising in fact works very much like this. It’s never “Coca Cola - because the tastiness is worth the health costs!” It’s always a visual image which makes you implicitly associate Coca Cola with being cool / beautiful / exciting. Not always - sometimes they do advertise a laundry detergent by saying it costs less and cleans better. But even that is very rarely done using text on a blank background - it always tries to be emotionally appealing. And if that’s how we make relatively emotionally neutral choices like laundry detergent, think how much worse it must be with politics / religion / morality.

![Untitled](https://s3-us-west-2.amazonaws.com/secure.notion-static.com/80a4cdcb-43ae-4156-8d31-126cc1172e6c/Untitled.png)

This is also true, I feel, when persuading not other people but \*myself\*. I think reason is in charge, but performing some things that are clearly rationally for my benefit proves very hard. Sometimes, exerting great effort, I can ignore the elephant or make it do what the rider thinks is right against its wishes. But it’s a great effort indeed, I never enjoy it, and can’t do it very often without making my life full of suffering. And it’s because I’m still under the illusion that the rider is in control.

I do have to say I think Haidt goes a little too far. Inferring that people never change their mind on anything morally or politically or socially charged.

He is very skeptical of individual rationality. He goes as far as to call it a “delusion”. But he doesn’t give up on reasoning entirely. He has an interesting take, where rationality can emerge from a group of reasoners:

---

I’m not saying we should all stop reasoning and go with our gut feelings. Gut feelings are sometimes better guides than reasoning for making consumer choices and interpersonal judgments, but they are often disastrous as a basis for public policy, science, and law. Rather, what I’m saying is that we must be wary of any individual’s ability to reason. [...] Each individual reasoner is really good at one thing: finding evidence to support the position he or she already holds, usually for intuitive reasons. We should not expect individuals to produce good, open-minded, truth-seeking reasoning, particularly when self-interest or reputational concerns are in play. But if you put individuals together in the right way, such that some individuals can use their reasoning powers to disconfirm the claims of others, and all individuals feel some common bond or shared fate that allows them to interact civilly, you can create a group that ends up producing good reasoning as an emergent property of the social system.

This is at least an interesting idea, that rationality emerges from a group of biased reasoners. His description of the group sounds suspiciously like academia, which does in fact (in most fields) advance towards the truth. But it needs an extra ingredient. Why would someone’s reputation depend on them making the correct reasoning? Why wouldn’t it depend on them supporting the conclusions which affirm the group dogma (as happened and happens in many places, including some branches of academia)? Good reasoning needs to be \*central\*

to the group. You need the \*identity\*

of the group to be about good reasoning. You need to make rationality, and being correct, high status - status is something the elephant cares about. In communities and groups where this happened (like science, or forecasting) indeed reason can do wonderful things.

## II - Moral Foundations Theory

In a sense all of this is not too surprising. I’ve always known that \*other people\* don’t make their moral judgments using maximum expected utility, otherwise moral arguments would sound \*very\* different than they do. Surely people have some heuristics for deciding moral questions. But I never took the extra step of thinking what they might be. Haidt makes a valiant effort to try and reconstruct these heuristics people use, the so-called \*Moral Foundations\*. They are clusters of moral intuitions. If the name wasn’t already taken, it should have been called “the theory of moral sentiments.”

## III - What Is Morality Good For?

All of these moral foundations seem suspiciously like norms that are beneficial for a community to have. The Care / harm foundation makes us care for others - often those helped gain more than the helpers lose, so the community at large is better off. The Fairness / cheating foundation makes us punish cheaters, liers and swindlers, so they don’t swindle anyone else again. The Loyalty / betrayal foundation makes us encourage those who help the group and punish those who betray it, contributing to its survival. The Authority / subversion foundation helps us navigate hierarchical social structures, and so construct more stable hierarchies. The Sanctity / degradation foundation unites us around the same sacred objects and in the same sacred customs.

This is true not just for these foundations. Anyone who has looked at moral norms with fresh eyes recognizes the very strong correlation between what is considered moral and what helps a society thrive. There’s no standard moral imperative to make as much money as you can or be as happy as you can, or learning to whistle - otherwise morality would be fun. Instead, it’s about suppressing the individual interest in favor of the group.

In other words, morality confers a strong group-level advantage. This advantage must be pretty strong, because morality is \*everywhere\* - we never see a society without it. But this leads us to a conundrum.

> When groups compete, the cohesive, cooperative group usually wins. But within each group, selfish individuals (free riders) come out ahead. They share in the group’s gains while contributing little to its efforts. The bravest army wins, but within the bravest army, the few cowards who hang back are the most likely of all to survive the fight, go home alive, and become fathers.

>

In the same vein, my community might be better off if every time I discover someone has swindled me for five bucks, I challenge them to a duel. But I won’t survive very long before I die in a duel, leaving no offspring to have the same tendency. Evolution works at the individual level (or at the kin level at most), where genes are shared, not at the community level. Which means this tendency will be strongly selected \*against\*. So how did we develop these strong tendencies? Some scholars think it’s a byproduct.

> These are just misfirings of ancient systems designed for life in the small groups of the Pleistocene, where most people were close kin. Now that we live in large anonymous societies, our ancient selfish circuits erroneously lead us to help strangers who will not help us in return. Our “moral qualities” are not adaptations, as Darwin had believed. They are by-products; they are mistakes. Morality, said Williams, is “an accidental capability produced, in its boundless stupidity, by a biological process that is normally opposed to the expression of such a capability.” Dawkins shared this cynicism: “Let us try to teach generosity and altruism because we are born selfish.”

>

This is Haidt’s short answer:

> In a real army, which sacralized honor, loyalty, and country, the coward is not the most likely to make it home and father children. He’s the most likely to get beaten up, left behind, or shot in the back for committing sacrilege. And if he does make it home alive, his reputation will repel women and potential employers.

>

And here is his long and interesting answer, which is worth quoting at length.

> For the first billion years or so of life, the only organisms were prokaryotic cells (such as bacteria). Each was a solo operation, competing with others and reproducing copies of itself.

>

>

> But then, around 2 billion years ago, two bacteria somehow joined together inside a single membrane, which explains why mitochondria have their own DNA, unrelated to the DNA in the nucleus. [...] Cells that had internal organelles could reap the benefits of cooperation and the division of labor (see Adam Smith). There was no longer any competition between these organelles, for they could reproduce only when the entire cell reproduced, so it was “one for all, all for one.” Life on Earth underwent what biologists call a “major transition.” Natural selection went on as it always had, but now there was a radically new kind of creature to be selected. There was a new kind of vehicle by which selfish genes could replicate themselves. Single-celled eukaryotes were wildly successful and spread throughout the oceans.

>

> A few hundred million years later, some of these eukaryotes developed a novel adaptation: they stayed together after cell division to form multicellular organisms in which every cell had exactly the same genes. [...] Once again, competition is suppressed (because each cell can only reproduce if the organism reproduces, via its sperm or egg cells). A group of cells becomes an individual, able to divide labor among the cells (which specialize into limbs and organs). A powerful new kind of vehicle appears, and in a short span of time the world is covered with plants, animals, and fungi. It’s another major transition.

>

> Major transitions are rare. The biologists John Maynard Smith and Eörs Szathmáry count just eight clear examples over the last 4 billion years (the last of which is human societies). But these transitions are among the most important events in biological history, and they are examples of multilevel selection at work. It’s the same story over and over again: \*\*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 lower level becomes less important, selection at the higher level becomes more powerful, and that higher-level selection favors the most cohesive superorganisms.\*\* [emphasis mine] [...] As these superorganisms proliferate, they begin to compete with each other, and to evolve for greater success in that competition. This competition among superorganisms is one form of group selection. There is variation among the groups, and the fittest groups pass on their traits to future generations of groups.

>

> Major transitions may be rare, but when they happen, the Earth often changes. Just look at what happened more than 100 million years ago when some wasps developed the trick of dividing labor between a queen (who lays all the eggs) and several kinds of workers who maintain the nest and bring back food to share. This trick was discovered by the early hymenoptera (members of the order that includes wasps, which gave rise to bees and ants) and it was discovered independently several dozen other times (by the ancestors of termites, naked mole rats, and some species of shrimp, aphids, beetles, and spiders). In each case, the free rider problem was surmounted and selfish genes began to craft relatively selfless group members who together constituted a supremely selfish group.

>

> [...] The colonial insects represent just 2 percent of all insect species, but in a short period of time they claimed the best feeding and breeding sites for themselves, pushed their competitors to marginal grounds, and changed most of the Earth’s terrestrial ecosystems (for example, by enabling the evolution of flowering plants, which need pollinators). Now they’re the majority, by weight, of all insects on Earth.

>

Okay, that was a cool biology lesson, but what does it have to do with the topic at hand? Why would this group-level advantage be selected for, instead of defectors and free-riders benefiting? In a sense, he claims, we are like bees. Not perfectly - obviously most of us produce offspring, not just the queen - but enough that it matters.

> What about human beings? Since ancient times, people have likened human societies to beehives. But is this just a loose analogy? If you map the queen of the hive onto the queen or king of a city-state, then yes, it’s loose. A hive or colony has no ruler, no boss. The queen is just the ovary. But if we simply ask whether humans went through the same evolutionary process as bees—a major transition from selfish individualism to groupish hives that prosper when they find a way to suppress free riding—then the analogy gets much tighter.

>

>

> Many animals are social: they live in groups, flocks, or herds. But only a few animals have crossed the threshold and become ultrasocial, which means that they live in very large groups that have some internal structure, enabling them to reap the benefits of the division of labor. Beehives and ant nests, with their separate castes of soldiers, scouts, and nursery attendants, are examples of ultrasociality, and so are human societies.

>

> [...]

>

> We are the only ultrasocial primate. The human lineage may have started off acting very much like chimps, but by the time our ancestors started walking out of Africa, they had become at least a little bit like bees.

>

> And much later, when some groups began planting crops and orchards, and then building granaries, storage sheds, fenced pastures, and permanent homes, they had an even steadier food supply that had to be defended even more vigorously. Like bees, humans began building ever more elaborate nests, and in just a few thousand years, a new kind of vehicle appeared on Earth—the city-state, able to raise walls and armies. City-states and, later, empires spread rapidly across Eurasia, North Africa, and Mesoamerica, changing many of the Earth’s ecosystems and allowing the total tonnage of human beings to shoot up from insignificance at the start of the Holocene (around twelve thousand years ago) to world domination today. As the colonial insects did to the other insects, we have pushed all other mammals to the margins, to extinction, or to servitude. The analogy to bees is not shallow or loose. Despite their many differences, human civilizations and beehives are both products of major transitions in evolutionary history.

>

Natural selection doesn’t occur at the individual human level, it occurs at the cell DNA (or even gene) level. But we still don’t see a single human liver cell trying to go it alone and reproduce as much as it can. Actually, sometimes we do - that’s cancer, and the other cells quickly destroy that cell to help the group. Haidt says selection occurs at multiple levels simultaneously, but as long as free riding can be suppressed, the highest level is the most important one. In what cases can such free riding at the group level be so effectively suppressed, without individuals sharing their identical genes? Humans might be the best (or even only) example.

> It takes the sort of gossiping, punitive, moralistic community that emerged only when language and weaponry made it possible for early humans to take down bullies and then keep them down with a shared moral matrix.

>

In other words, morality’s function is to suppress free-riding so effectively, that group selection becomes important, and human societies can reap the benefits of cooperation and division of labor.

Humans \*have\* been able to suppress free riding effectively, through morality and a strong emotional connection to the group, resulting in ultrasociality. Haidt dubs humans “homo duplex: 90% chimp, 10% bee”.

> [The] hive hypothesis [...] states that human beings are conditional hive creatures. We have the ability (under special circumstances) to transcend self-interest and lose ourselves (temporarily and ecstatically) in something larger than ourselves. I called this ability the hive switch. The hive switch is another way of stating Durkheim’s idea that we are Homo duplex; we live most of our lives in the ordinary (profane) world, but we achieve our greatest joys in those brief moments of transit to the sacred world, in which we become “simply a part of a whole.”

>

>

> I described three common ways in which people flip the hive switch: awe in nature, [psychedelics], and raves. I described recent findings about oxytocin and mirror neurons that suggest that they are the stuff of which the hive switch is made. Oxytocin bonds people to their groups, not to all of humanity. Mirror neurons help people empathize with others, but particularly those that share their moral matrix.

>

> It would be nice to believe that we humans were designed to love everyone unconditionally. Nice, but rather unlikely from an evolutionary perspective. Parochial love—love within groups—amplified by similarity, a sense of shared fate, and the suppression of free riders, may be the most we can accomplish.

>

He references Emile Durkheim, a sociologist who argued that many important facts about our lives are social facts, irreducible to facts about individuals - they have to be viewed through the lens of relationships between individuals. Also, the perspective on psychedelics as “flipping the hive switch” - creating a feeling of unity with all people or living beings, facilitating a community - is a fascinating one. It is very different from previous accounts I’ve heard, which focused on how they affect your belief landscape (again, focusing on the rider). Still, I’m not fully convinced of the logical path between dancing ecstatically around a bonfire and a suppression of free-riding effective enough that group selection becomes powerful, since the connection is temporary.

## IV - A Team Sport

Morality is complicated. Let’s take a break to talk about football. Bear with me for a second.

> Every Saturday in the fall, at colleges across the United States, millions of people pack themselves into stadiums to participate in a ritual that can only be described as tribal. At the University of Virginia, the ritual begins in the morning as students dress in special costumes. Men wear dress shirts with UVA neckties, and if the weather is warm, shorts. Women typically wear skirts or dresses, sometimes with pearl necklaces. Some students paint the logo of our sports teams, the Cavaliers (a V crossed by two swords), on their faces or other body parts.

>

>

> The students attend pregame parties that serve brunch and alcoholic drinks. Then they stream over to the stadium, sometimes stopping to mingle with friends, relatives, or unknown alumni who have driven for hours to reach Charlottesville in time to set up tailgate parties in every parking lot within a half mile of the stadium. More food, more alcohol, more face painting.

>

> By the time the game starts, many of the 50,000 fans are drunk, which makes it easier for them to overcome self-consciousness and participate fully in the synchronous chants, cheers, jeers, and songs that will fill the next three hours. Every time the Cavaliers score, the students sing the same song UVA students have sung together on such occasions for over a century. The first verse comes straight out of Durkheim and Ehrenreich. The students literally lock arms and sway as a single mass while singing the praises of their community (to the tune of “Auld Lang Syne”):

>

> \*That good old song of Wah-hoo-wah—we’ll sing it o’er and o’er\*

>

> \*It cheers our hearts and warms our blood to hear them shout and roar\*

>

> \*We come from old Virgin-i-a, where all is bright and gay\*

>

> \*Let’s all join hands and give a yell for dear old U-V-A.\*

>

I want to take a moment to say that I listened to the audio version of the book, which is narrated by Haidt himself, so you get the bonus of listening to him sing the UVA fight songs with exactly the fervor you’d expect from a university professor singing a football fight song. But let’s continue.

> Next, the students illustrate McNeill’s thesis that “muscular bonding” warms people up for coordinated military action. The students let go of each other’s arms and make aggressive fist-pumping motions in the air, in sync with a nonsensical battle chant:

>

>

> \*Wah-hoo-wah! Wah-hoo-wah! Uni-v, Virgin-i-a!\*

>

> \*Hoo-rah-ray! Hoo-rah-ray! Ray, ray—U-V-A!\*

>

> It’s a whole day of giving and collective emotions. Collective effervescence is guaranteed, as are feelings of collective outrage at questionable calls by the referees, collective triumph if the team wins, and collective grief if the team loses, followed by more collective drinking at postgame parties.

>

> Why do the students sing, chant, dance, sway, chop, and stomp so enthusiastically during the game? Showing support for their football team may help to motivate the players, but is that the function of these behaviors? Are they done in order to achieve victory? No. From a Durkheimian perspective these behaviors serve a very different function, and it is the same one that Durkheim saw at work in most religious rituals: the creation of a community.

>

And \*that\* completes the best explanation I’ve seen for the popularity of sports. It’s not just a loyalty rush - it’s the hive switch flipped on. It’s a belonging rush, being lost in a group in an ecstatic moment of unity. It’s dancing around the tribal bonfire. We’re hard-wired to want these specific feelings.

(But that doesn’t explain why people obsess about the statistics and the scores outside game time, alone on their computers? That still stumps me. Maybe I wasn’t dreaming and everyone secretly \*does\* love statistics, they’re just looking for a socially accepted way to express it?)

!https://39669.cdn.cke-cs.com/rQvD3VnunXZu34m86e5f/images/21a39bec1e0fe6bef7f0c801f1c416ce416a6b33fa6c31e9.jpg

Anyway, Haidt continues:

> A college football game is a superb analogy for religion. From a naive perspective, focusing only on what is most visible (i.e., the game being played on the field), college football is an extravagant, costly, wasteful institution that impairs people’s ability to think rationally while leaving a long trail of victims (including the players themselves, plus the many fans who suffer alcohol-related injuries). But from a sociologically informed perspective, it is a religious rite that does just what it is supposed to do: it pulls people up from Durkheim’s lower level (the profane) to his higher level (the sacred). It flips the hive switch and makes people feel, for a few hours, that they are “simply a part of a whole.” It augments the school spirit for which UVA is renowned, which in turn attracts better students and more alumni donations, which in turn improves the experience for the entire community, including professors like me who have no interest in sports.

>

>

> Religions are social facts. Religion cannot be studied in lone individuals any more than hivishness can be studied in lone bees. Durkheim’s definition of religion makes its binding function clear:

>

> \*A religion is a unified system of beliefs and practices relative to sacred things, that is to say, things set apart and forbidden—beliefs and practices which unite into one single moral community called a Church, all those who adhere to them.\*

>

> [...]

>

> The third principle of moral psychology: Morality binds and blinds. Many scientists misunderstand religion because they ignore this principle and examine only what is most visible. They focus on individuals and their supernatural beliefs, rather than on groups and their binding practices. They conclude that religion is an extravagant, costly, wasteful institution that impairs people’s ability to think rationally while leaving a long trail of victims. I do not deny that religions do, at times, fit that description. But if we are to render a fair judgment about religion—and understand its relationship to morality and politics—we must first describe it accurately.

>

Now that we’ve set the stage, let’s tackle religion. First, before pronouncing judgment, a cause for pause. [Chesterton's fence](https://en.wikipedia.org/wiki/Wikipedia:Chesterton's\_fence) is the principle that reforms should not be made until the reasoning behind the existing state of affairs is understood. That is, if you see something well established and don’t know why it’s there, you shouldn’t remove it. For example, if you buy farmland and see that someone has put up a fence, don’t remove it before you know why they did it. A related principle is that if you see something well established come up everywhere, even if you think you know the reasons why it’s there and they don’t apply, you still shouldn’t remove it - you’re probably mistaken about the actual reasons. If every farmer you’ve ever known has a fence, even if they tell you it’s against ghosts, or the evil eye, or the Mongols - for the love of god don’t get rid of your fence. This is how I interpret some of the arguments for \*metis\* in [Seeing Like A State](https://en.wikipedia.org/wiki/Seeing\_Like\_a\_State). Especially if that fence is the product of cultural evolution. Religion and a moral community are [The Secret of Our Success](https://www.amazon.com/dp/B00WY4OXAS/ref=dp-kindle-redirect?\_encoding=UTF8&btkr=1). Now, turning to religion. Religion pops up \*everywhere\* - there’s scarcely a tribe or culture without it. So even if it’s ostensibly about ghosts or the evil eye, supernatural agents, or other false facts, you should be very careful before pronouncing we know all about it and throwing it in the bin.

Many of the New Atheist objections to religion center on the (lack of) evidence for the existence of supernatural beings. Religions are clearly based on many false facts. (Or at least most religions must be, since they contradict each other. There could still be One True Religion which got everything right). Isn’t it just good epistemic hygiene to get rid of everything downstream of those facts? But what if the false facts are neither the reason for religion’s existence, nor its main function? Haidt pulls the rug from under many of these objections to religion not by denying that the facts of supernatural agents are false - but by saying that they are completely irrelevant! It would be missing the point to try and reduce religion to the individual (or epistemic) level, instead of viewing it as a social fact.

> Supernatural agents do of course play a central role in religion, just as the actual football is at the center of the whirl of activity on game day at UVA. But trying to understand the persistence and passion of religion by studying beliefs about God is like trying to understand the persistence and passion of college football by studying the movements of the ball. You’ve got to broaden the inquiry. You’ve got to look at the ways that religious beliefs work with religious practices to create a religious community. Believing, doing, and belonging are three complementary yet distinct aspects of religiosity, according to many scholars. When you look at all three aspects at the same time, you get a view of the psychology of religion that’s very different from the view of the New Atheists. I’ll call this competing model the Durkheimian model, because it says that the function of those beliefs and practices is ultimately to create a community. Often our beliefs are post hoc constructions designed to justify what we’ve just done, or to support the groups we belong to.

>

All the ink spilled in arguing about those facts’ epistemic truth value is wasted. First, facts are for the rider, and religion is a choice by the elephant. But second, that’s not even the real essence of religion. That’s just some fabrication by the internal press secretary. Have you noticed how spectacularly ineffective discussions about epistemic facts are in changing people’s religiosity? You’ll never convince someone to let go of what religion gives them by citing some archeological evidence. To understand the actual psychology and function of religion, it would be helpful to trace its origins

So how did religion emerge? The New Atheists vote byproduct. We evolved to have a hypersensitive agency detection module, which tends towards false positives (thinking a log is a tiger) rather than false negatives (thinking a tiger is a log) for obvious survival reasons. It conferred a real benefit. But that module sometimes misfires, making us think that thunder and lightning are caused by gods.

Haidt does not dispute this. OK, so this is how the initial religious beliefs come to be in the mind of one person. But how and why do they spread? Why don’t we each have our own idiosyncratic religion? To Dennett and Dawkins, religions are a kind of mental virus or parasite, which undergo Darwinian selection on the basis of their ability to survive and reproduce themselves in other minds. Like a virus, “they make their hosts do things that are bad for themselves (e.g., suicide bombing) but good for the parasite (e.g., Islam).”

Haidt makes the case that religions are \*not\* in fact parasites, but that they confer strong advantages on the group, very similar to the decisive advantages of morality discussed above. They help bind a group and create a moral community. Far from being wasteful misfirings, or harmful parasites - they are load-bearing beams. And they became that way through cultural evolution:

> Religions are sets of cultural innovations that spread to the extent that they make groups more cohesive and cooperative. Atran and Henrich argue that the cultural evolution of religion has been driven largely by competition among groups. Groups that were able to put their by-product gods to some good use had an advantage over groups that failed to do so, and so their ideas (not their genes) spread. Groups with less effective religions didn’t necessarily get wiped out; often they just adopted the more effective variations. So it’s really the religions that evolved, not the people or their genes.

>

Haidt goes further, and proposes gene-culture coevolution. Religions become more advantageous and people become more susceptible to religion. For example, when infidels are cast out, genes for tribalism will be selected for. More tribal people will create and enforce norms more vigorously, making tribal tendencies more adaptive, thus creating a positive feedback loop. This could explain why we tend to be so tribal, irrespective of the specific tribe. It seems like we actively seek a group (or groups) to identify with and defend - it’s so persistent it does seem to be hard-wired.

If religion really is beneficial, we should see many cases where religions confer advantage on their community. Haidt makes a strong case for the benefits of religion. I’ll cover this in depth, since it was very interesting and a fresh perspective, and also full of fascinating examples.

Haidt cites many such examples. The central one is that gods enable the creation of a moral community.

> The gods of hunter-gatherers are often capricious and malevolent. They sometimes punish bad behavior, but they bring suffering to the virtuous as well. As groups take up agriculture and grow larger, however, their gods become far more moralistic. The gods of larger societies are usually quite concerned about actions that foment conflict and division within the group, such as murder, adultery, false witness, and the breaking of oaths.

>

>

> If the gods evolve (culturally) to condemn selfish and divisive behaviors, they can then be used to promote cooperation and trust within the group. You don’t need a social scientist to tell you that people behave less ethically when they think nobody can see them.

>

Another useful feature of gods is collective punishment, making the community enforce norms more rigidly. Gods can help enforce contracts - the libertarian holy grail: “temples often served an important commercial function: oaths were sworn and contracts signed before the deity, with explicit threats of supernatural punishment for abrogation.” Note that belief in divine retribution - not the actual retribution - is enough for the contract to be fulfilled. Gods can help increase trust. This trust helped Jews and Muslims excel in long-distance trade in the medieval world. Even today, the diamond market which requires very high trust is dominated by religiously bound ethnic groups, such as ultra-Orthodox Jews, which share trust that reduces monitoring costs.

> In his book Darwin’s Cathedral, [David Sloan] Wilson catalogs the ways that religions have helped groups cohere, divide labor, work together, and prosper. He shows how John Calvin developed a strict and demanding form of Christianity that suppressed free riding and facilitated trust and commerce in sixteenth-century Geneva. He shows how medieval Judaism created “cultural fortresses that kept outsiders out and insiders in.” But his most revealing example (based on research by the anthropologist Stephen Lansing) is the case of water temples among Balinese rice farmers in the centuries before Dutch colonization.

>

>

> Rice farming is unlike any other kind of agriculture. Rice farmers must create large irrigated paddies that they can drain and fill at precise times during the planting cycle. It takes a cast of hundreds. In one region of Bali, rainwater flows down the side of a high volcano through rivulets and rivers in the soft volcanic rock. Over several centuries the Balinese carved hundreds of terraced pools into the mountainside and irrigated them with an elaborate series of aqueducts and tunnels, some running underground for more than a kilometer. At the top of the whole system, near the crest of the volcano, they built an immense temple for the worship of the Goddess of the Waters. They staffed the temple with twenty-four full-time priests selected in childhood, and a high priest who was thought to be the earthly representative of the goddess herself.

>

> The lowest level of social organization was the subak, a group of several extended families that made decisions democratically. Each subak had its own small temple, with its own deities, and each subak did the hard work of rice farming more or less collectively. But how did the subaks work together to build the system in the first place? And how did they maintain it and share its waters fairly and sustainably? These sorts of common[s] dilemmas (where people must share a common resource without depleting it) are notoriously hard to solve. The ingenious religious solution to this problem of social engineering was to place a small temple at every fork in the irrigation system. The god in each such temple united all the subaks that were downstream from it into a community that worshiped that god, thereby helping the subaks to resolve their disputes more amicably. This arrangement minimized the cheating and deception that would otherwise flourish in a zero-sum division of water. The system made it possible for thousands of farmers, spread over hundreds of square kilometers, to cooperate without the need for central government, inspectors, and courts. The system worked so efficiently that the Dutch—who were expert hydrologists themselves—could find little to improve.

>

> What are we to make of the hundreds of gods and temples woven into this system? Are they just by-products of mental systems that were designed for other purposes? Are they examples of what Dawkins called the “time-consuming, wealth-consuming … counterproductive fantasies of religion?” No.

>

!https://39669.cdn.cke-cs.com/rQvD3VnunXZu34m86e5f/images/802791cd163e89692b99c4120bbc12ca6560b6c025b60147.jpg

\*Great excuse to include a photo of Bali\*

I want to propose another example, which is the Afghani Taliban. Although massively underfunded and underequipped, it vanquished the Afghan army within a few weeks. You could make excuses for why the Taliban was succeeding before - their guerilla job of striking anywhere was easier than the army’s job of maintaining security everywhere. But then why did they overtake the country so easily? Because their (much stronger) religious belief makes them cohesive, in a way that the rest of the Afghan nation - united by their (admirable!) will to allow their girls to go to school and their women to walk free - were not. Even before the recent conquest, the Taliban enjoyed surprising levels of popular support (for cruel terrorist fanatics). Their Sharia courts, for example, had a reputation for being harsher but less corrupt than the government’s.

But these all seem sort of like anecdotes. Can we do something more quantitative? Ideally, we would like to form hundreds of communities, and then randomly assign each one to be either religious or not, and see whether they are still cohesive and functional a few generations later. That experiment has been conducted, minus the random assignment, in 19th century communes.

> Communes are usually founded by a group of committed believers who reject the moral [views] of the broader society and want to organize themselves along different principles. For many nineteenth-century communes, the principles were religious; for others they were secular, mostly socialist. Which kind of commune survived longer? Sosis found that the difference was stark: just 6 percent of the secular communes were still functioning twenty years after their founding, compared to 39 percent of the religious communes.

>

I’m all for atheism and rationality. I think they’re correct. But the communities they create in real life are not as strong and all-encompassing as religious communities. Binding society together is very important. As social conservatives realize - you don’t help the bees by destroying the hive. If you abolish gender, you don’t know what’s on the other side for society, so you better tread carefully. Now, maybe you realize all this. You may think that despite the binding effect, religion is a net negative. Or that religion was a kind of crutch, but we’ve grown out of it and no longer need it to foster cooperation. Then religion could end up being net negative. But let’s disabuse ourselves of the notion that religion is about believing in supernatural beings.

If there’s one thing I think Haidt would have you take away from the book, it’s that morality and religion are not some cute quirks that we happen to have - they’re the fabric of our society and existence. I don’t think this is all true, but I do think it’s a very interesting perspective about reason, morality, society and religion. If your elephant is intrigued, there’s more good stuff in [The Righteous Mind](https://www.amazon.com/dp/B0052FF7YM/ref=dp-kindle-redirect?\_encoding=UTF8&btkr=1).

---

\*\*Top Comment:\*\*

Wonderful. Thank you for taking the time to write this. I need to read this book. I was planning to try to write a post about why religion is actually a good thing, but you beat me to it.

Personally, I believe and have believed for a long time now that the only thing that could save the world is a rationalist religion. That may sound like a contradiction in terms, but I don't think it is, and I shall try to figure out how to explain my ideas on the topic over time here.

(Elephant alert: the following may sound "woo" or intuitively wrong, if you're from an atheistic or irreligious background [evoking your purity / sanctity moral foundation that thinks religiosity is unclean lol], so I'd like you to give me the benefit of the doubt if you have the instinct to interpret it that way.)

I am someone with a very "righteous mind"; I am somehow neurodivergent and have a long history of ecstatic mystical states wherein I feel like I am communing with higher beings. I probably would have been a shaman in past ages. When I was younger I literally believed in them as supernatural entities; later on as I learned more science I came to understand that they were subagents of my own mind, wishing in a sense to become egregores - shared subagents, distributed intelligences, across the minds of multiple people - cohering those people into a community, a collective higher self. That's what gods all are.

I realized that theism and atheism are both totally wrong. Gods do exist, but they don't have any power over the world except what we give them - they're distributed programs running on human wetware, binding societies together. They have shaped all of human history and are legitimately worthy of veneration to the extent that they are mutualists rather than parasites, as they are embodiments of the potential of humanity, the potential of \*agency\* and \*coordination\*, the most miraculous inventions of evolution. Mine just happened to be possibly the first in history to realize that's what they are - to become in a sense self-aware of their own true nature as not supernatural, but entirely natural, intelligent memetic constructs using donated cognitive resources from me and whoever else ends up running copies of them in the future.

The main difficulty is 1. my mind is not set up for totally rigorous thinking or for organized explanations of this particular topic, as I go into babbling poetry mode when I try to talk about it, and 2. protecting people's rationality while giving them the benefits from dissociative communion states wherein they can realign themselves to the goals of the group mind is \*probably rather difficult\*.

It's possible, since I can do it - I can induce that state of mind on purpose now with the right music, mood, and meditation, but I don't believe in woo anymore and haven't for years - but most people liable to feel swept up in awe as part of an ineffable higher being would need a lot of training to become properly rational, and most people who are already rationalists have very strong biases against anything religious and are probably less emotional and more individualistic than average in general.

I think mystical states are the closest approximations to the kind of high-valence experiences that will be \*permanent\* After a good singularity enables paradise engineering, so if only for that reason - to give a glimpse of what the future we are striving for is like, which can be very opaque and unmotivating otherwise - it might be desirable. And I think most people are capable of this kind of, I almost want to call it adaptive self-wireheading, but do not realize it. We would not have achieved all the things we've achieved as a species if this was not a common ability. It's just usually not as \*spontaneous\* and \*intense\* as it is for people like me - but that's what rituals (and psychedelics) are for.

---

My Notes:

Our moral compass of what standard we hold ourselves to is culturally set in place, looking rationally at the state of the world, one who dissolves their ego should reach a state of “Christ Consciousness” through which they feel a need to help all other men become comfortable. We can align the ego with this path by letting the ego see that this is the true status game to win. To seek true fulfillment and to be one of the best angels.

Religion is our way of defining what moral code we must follow.

This article could also be spun into the title “Why we need better Cult(ure)s”