Jonathan Haidt  
Moral Psychology and the Misunderstanding of Religion

I study morality from every angle I can find. Morality is one of those basic aspects of humanity, like sexuality and eating, that can’t fit into one or two academic fields. I think morality is unique, however, in having a kind of spell that disguises it. We all care about morality so passionately that it’s hard to look straight at it. We all look at the world through some kind of moral lens, and because most of the academic community uses the same lens, we validate each other’s visions and distortions. I think this problem is particularly acute in some of the new scientific writing about religion.

When I started graduate school at Penn in 1987, it seemed that developmental psychology owned the rights to morality within psychology. Everyone was either using or critiquing Lawrence Kohlberg’s ideas, as well as his general method of interviewing kids about dilemmas (such as: should Heinz steal a drug to save his wife’s life?). Everyone was studying how children’s understanding of moral concepts changed with experience. But in the 1990s two books were published that I believe triggered an explosion of cross-disciplinary scientific interest in morality, out of which has come a new synthesis—very much along the lines that E. O. Wilson predicted in 1975.

The first was Antonio Damasio’s *Descartes’ Error*, in 1994, which showed a very broad audience that morality could be studied using the then new technology of fMRI, and also that morality, and rationality itself, were crucially dependent on the proper functioning of emotional circuits in the prefrontal cortex. The second was Frans de Waal’s *Good Natured*, published just two years later, which showed an equally broad audience that the building blocks of human morality are found in other apes and are products of natural selection in the highly social primate lineage. These two books came out just as John Bargh was showing social psychologists that automatic and unconscious processes can and probably do cause the majority of our behaviors, even morally loaded actions (like rudeness or altruism) that we thought we were controlling consciously.

Furthermore, Damasio and Bargh both found, as Michael Gazzaniga had years before, that people couldn’t stop themselves from making up post-hoc explanations for whatever it was they had just done for unconscious reasons. Combine these developments and suddenly Kohlbergian moral psychology seemed to be studying the wagging tail, rather than the dog. If the building blocks of morality were shaped by natural selection long before language arose, and if those evolved structures work largely by giving us feelings that shape our behavior automatically, then why should we be focusing on the verbal reasons that people give to explain their judgments in hypothetical moral dilemmas?

In my dissertation and my other early studies, I told people short stories in which a person does something disgusting or disrespectful that was perfectly harmless (for example, a family cooks and eats its dog, after the dog was killed by a car). I was trying to pit the emotion of disgust against reasoning about harm and individual rights.

I found that disgust won in nearly all groups I studied (in Brazil, India, and the United States), except for groups of politically liberal college students, particularly Americans, who overrode their disgust and said that people have a right to do whatever they want, as long as they don’t hurt anyone else.
These findings suggested that emotion played a bigger role than the cognitive developmentalists had given it. These findings also suggested that there were important cultural differences, and that academic researchers may have inappropriately focused on reasoning about harm and rights because we primarily study people like ourselves—college students, and also children in private schools near our universities, whose morality is not representative of the United States, let alone the world.

So in the 1990s I was thinking about the role of emotion in moral judgment, I was reading Damasio, De Waal, and Bargh, and I was getting very excited by the synergy and consilience across disciplines. I wrote a review article called "The Emotional Dog and its Rational Tail," which was published in 2001, a month after Josh Greene's enormously influential Science article. Greene used fMRI to show that emotional responses in the brain, not abstract principles of philosophy, explain why people think various forms of the "trolley problem" (in which you have to choose between killing one person or letting five die) are morally different.

Obviously I'm biased in terms of what I notice, but it seems to me that the zeitgeist in moral psychology has changed since 2001. Most people who study morality now read and write about emotions, the brain, chimpanzees, and evolution, as well as reasoning. This is exactly what E. O. Wilson predicted in Sociobiology: that the old approaches to morality, including Kohlberg's, would be swept away or merged into a new approach that focused on the emotive centers of the brain as biological adaptations. Wilson even said that these emotive centers give us moral intuitions, which the moral philosophers then justify while pretending that they are intuiting truths that are independent of the contingencies of our evolved minds.

And now, 30 years later, Josh Greene has a paper in press where he uses neuroscientific evidence to reinterpret Kantian deontological philosophy as a sophisticated post-hoc justification of our gut feelings about rights and respect for other individuals. I think E. O. Wilson deserves more credit than he gets for seeing into the real nature of morality and for predicting the future of moral psychology so uncannily. He's in my pantheon, along with David Hume and Charles Darwin. All three were visionaries who urged us to focus on the moral emotions and their social utility.

I recently summarized this new synthesis in moral psychology with four principles:

1) Intuitive primacy but not dictatorship. This is the idea, going back to Wilhelm Wundt and channeled through Robert Zajonc and John Bargh, that the mind is driven by constant flashes of affect in response to everything we see and hear.

Our brains, like other animal brains, are constantly trying to fine tune and speed up the central decision of all action: approach or avoid. You can't understand the river of fMRI studies on neuroeconomics and decision making without embracing this principle. We have affectively-valenced intuitive reactions to almost everything, particularly to morally relevant stimuli such as gossip or the evening news. Reasoning by its very nature is slow, playing out in seconds.

Studies of everyday reasoning show that we usually use reason to search for evidence to support our initial judgment, which was made in milliseconds. But I do agree with Josh Greene that sometimes we can use controlled processes such as reasoning to override our initial intuitions. I just think this happens rarely, maybe in one or two percent of the hundreds of judgments we make each week. And I do agree with Marc Hauser that these moral intuitions require a lot of computation, which he is unpacking.
Hauser and I mostly disagree on a definitional question: whether this means that "cognition" precedes "emotion." I try never to contrast those terms, because it's all cognition. I think the crucial contrast is between two kinds of cognition: intuitions (which are fast and usually affectively laden) and reasoning (which is slow, cool, and less motivating).

2) **Moral thinking is for social doing.** This is a play on William James' pragmatist dictum that thinking is for doing, updated by newer work on Machiavellian intelligence. The basic idea is that we did not evolve language and reasoning because they helped us to find truth; we evolved these skills because they were useful to their bearers, and among their greatest benefits were reputation management and manipulation.

Just look at your stream of consciousness when you are thinking about a politician you dislike, or when you have just had a minor disagreement with your spouse. It's like you're preparing for a court appearance. Your reasoning abilities are pressed into service generating arguments to defend your side and attack the other. We are certainly able to reason dispassionately when we have no gut feeling about a case, and no stake in its outcome, but with moral disagreements that's rarely the case. As David Hume said long ago, reason is the servant of the passions.

3) **Morality binds and builds.** This is the idea stated most forcefully by Emile Durkheim that morality is a set of constraints that binds people together into an emergent collective entity.

Durkheim focused on the benefits that accrue to individuals from being tied in and restrained by a moral order. In his book *Suicide* he alerted us to the ways that freedom and wealth almost inevitably foster anomie, the dangerous state where norms are unclear and people feel that they can do whatever they want.

Durkheim didn't talk much about conflict between groups, but Darwin thought that such conflicts may have spurred the evolution of human morality. Virtues that bind people to other members of the tribe and encourage self-sacrifice would lead virtuous tribes to vanquish more selfish ones, which would make these traits more prevalent.

Of course, this simple analysis falls prey to the free-rider problem that George Williams and Richard Dawkins wrote so persuasively about. But I think the terms of this debate over group selection have changed radically in the last 10 years, as culture and religion have become central to discussions of the evolution of morality.

I'll say more about group selection in a moment. For now I just want to make the point that humans do form tight, cooperative groups that pursue collective ends and punish cheaters and slackers, and they do this most strongly when in conflict with other groups. Morality is what makes all of that possible.

4) **Morality is about more than harm and fairness.** In moral psychology and moral philosophy, morality is almost always about how people treat each other. Here's an influential definition from the Berkeley psychologist Elliot Turiel: morality refers to "prescriptive judgments of justice, rights, and welfare pertaining to how people ought to relate to each other."
Kohlberg thought that all of morality, including concerns about the welfare of others, could be derived from the psychology of justice. Carol Gilligan convinced the field that an ethic of "care" had a separate developmental trajectory, and was not derived from concerns about justice.

OK, so there are two psychological systems, one about fairness/justice, and one about care and protection of the vulnerable. And if you look at the many books on the evolution of morality, most of them focus exclusively on those two systems, with long discussions of Robert Trivers' reciprocal altruism (to explain fairness) and of kin altruism and/or attachment theory to explain why we don't like to see suffering and often care for people who are not our children.

But if you try to apply this two-foundation morality to the rest of the world, you either fail or you become Procrustes. Most traditional societies care about a lot more than harm/care and fairness/justice. Why do so many societies care deeply and morally about menstruation, food taboos, sexuality, and respect for elders and the Gods? You can't just dismiss this stuff as social convention. If you want to describe human morality, rather than the morality of educated Western academics, you've got to include the Durkheimian view that morality is in large part about binding people together.

From a review of the anthropological and evolutionary literatures, Craig Joseph (at Northwestern University) and I concluded that there were three best candidates for being additional psychological foundations of morality, beyond harm/care and fairness/justice. These three we label *ingroup/loyalty* (which may have evolved from the long history of cross-group or sub-group competition, related to what Joe Henrich calls "coalitional psychology"); *authority/respect* (which may have evolved from the long history of primate hierarchy, modified by cultural limitations on power and bullying, as documented by Christopher Boehm), and *purity/sanctity*, which may be a much more recent system, growing out of the uniquely human emotion of disgust, which seems to give people feelings that some ways of living and acting are higher, more noble, and less carnal than others.

Joseph and I think of these foundational systems as expressions of what Dan Sperber calls "learning modules"—they are evolved modular systems that generate, during enculturation, large numbers of more specific modules which help children recognize, quickly and automatically, examples of culturally emphasized virtues and vices. For example, we academics have extremely fine-tuned receptors for sexism (related to fairness) but not sacrilege (related to purity).

Virtues are socially constructed and socially learned, but these processes are highly prepared and constrained by the evolved mind. We call these three additional foundations the *binding* foundations, because the virtues, practices, and institutions they generate function to bind people together into hierarchically organized interdependent social groups that try to regulate the daily lives and personal habits of their members. We contrast these to the two *individualizing* foundations (harm/care and fairness/reciprocity), which generate virtues and practices that protect individuals from each other and allow them to live in harmony as autonomous agents who can focus on their own goals.

My UVA colleagues Jesse Graham, Brian Nosek, and I have collected data from about 7,000 people so far on a survey designed to measure people's endorsement of these five foundations. In every sample we've looked at, in the United States and in other Western countries, we find that people who self-identify as liberals endorse moral values and statements related to the two individualizing foundations primarily, whereas self-described conservatives endorse values and statements related
to all five foundations. It seems that the moral domain encompasses more for conservatives—it’s not just about Gilligan’s care and Kohlberg’s justice. It’s also about Durkheim’s issues of loyalty to the group, respect for authority, and sacredness.

I hope you’ll accept that as a purely descriptive statement. You can still reject the three binding foundations normatively—that is, you can still insist that ingroup, authority, and purity refer to ancient and dangerous psychological systems that underlie fascism, racism, and homophobia, and you can still claim that liberals are right to reject those foundations and build their moral systems using primarily the harm/care and fairness/reciprocity foundations.

But just go with me for a moment that there is this difference, descriptively, between the moral worlds of secular liberals on the one hand and religious conservatives on the other. There are, of course, many other groups, such as the religious left and the libertarian right, but I think it’s fair to say that the major players in the new religion wars are secular liberals criticizing religious conservatives. Because the conflict is a moral conflict, we should be able to apply the four principles of the new synthesis in moral psychology.

In what follows I will take it for granted that religion is a part of the natural world that is appropriately studied by the methods of science. Whether or not God exists (and as an atheist I personally doubt it), religiosity is an enormously important fact about our species. There must be some combination of evolutionary, developmental, neuropsychological, and anthropological theories that can explain why human religious practices take the various forms that they do, many of which are so similar across cultures and eras. I will also take it for granted that religious fundamentalists, and most of those who argue for the existence of God, illustrate the first three principles of moral psychology (intuitive primacy, post-hoc reasoning guided by utility, and a strong sense of belonging to a group bound together by shared moral commitments).

But because the new atheists talk so much about the virtues of science and our shared commitment to reason and evidence, I think it’s appropriate to hold them to a higher standard than their opponents. Do these new atheist books model the scientific mind at its best? Or do they reveal normal human beings acting on the basis of their normal moral psychology?

1) **Intuitive primacy but not dictatorship.** It’s clear that Richard Dawkins (in *The God Delusion*) and Sam Harris (in *Letter To A Christian Nation*) have strong feelings about religion in general and religious fundamentalists in particular. Given the hate mail they receive, I don’t blame them. The passions of Dawkins and Harris don’t mean that they are wrong, or that they can’t be trusted. One can certainly do good scholarship on slavery while hating slavery.

But the presence of passions should alert us that the authors, being human, are likely to have great difficulty searching for and then fairly evaluating evidence that opposes their intuitive feelings about religion. We can turn to Dawkins and Harris to make the case for the prosecution, which they do brilliantly, but if we readers are to judge religion we will have to find a defense attorney. Or at least we’ll have to let the accused speak.

2) **Moral thinking is for social doing.** This is where the scientific mind is supposed to depart from the lay mind. The normal person (once animated by emotion) engages in moral reasoning to find ammunition, not truth; the normal person attacks the motives and character of her opponents when it will be advantageous to do so. The scientist, in contrast, respects empirical evidence as the ultimate authority and avoids *ad hominem* arguments. The metaphor for science is a voyage of
discovery, not a war. Yet when I read the new atheist books, I see few new shores. Instead I see battlefields strewn with the corpses of straw men. To name three:

a) The new atheists treat religions as sets of beliefs about the world, many of which are demonstrably false. Yet anthropologists and sociologists who study religion stress the role of ritual and community much more than of factual beliefs about the creation of the world or life after death.

b) The new atheists assume that believers, particularly fundamentalists, take their sacred texts literally. Yet ethnographies of fundamentalist communities (such as James Ault’s *Spirit and Flesh*) show that even when people claim to be biblical literalists, they are in fact quite flexible, drawing on the bible selectively—or ignoring it—to justify humane and often quite modern responses to complex social situations.

c) The new atheists all review recent research on religion and conclude that it is an evolutionary byproduct, not an adaptation. They compare religious sentiments to moths flying into candle flames, ants whose brains have been hijacked for a parasite’s benefit, and cold viruses that are universal in human societies. This denial of adaptation is helpful for their argument that religion is bad for people, even when people think otherwise.

I quite agree with these authors’ praise of the work of Pascal Boyer and Scott Atran, who have shown how belief in supernatural entities may indeed be an accidental output of cognitive systems that otherwise do a good job of identifying objects and agents. Yet even if belief in gods was initially a byproduct, as long as such beliefs had consequences for behavior then it seems likely that natural selection operated upon phenotypic variation and favored the success of individuals and groups that found ways (genetic or cultural or both) to use these gods to their advantage, for example as commitment devices that enhanced cooperation, trust, and mutual aid.

3) **Morality binds and builds.** Dawkins is explicit that his goal is to start a movement, to raise consciousness, and to arm atheists with the arguments they’ll need to do battle with believers. The view that “we” are virtuous and our opponents are evil is a crucial step in uniting people behind a cause, and there is plenty of that in the new atheist books. A second crucial step is to identify traitors in our midst and punish or humiliate them. There is some of that too in these books—atheists who defend the utility of religion or who argue for disengagement or détente between science and religion are compared to Chamberlain and his appeasement of Hitler.

To my mind an irony of Dawkins’ position is that he reveals a kind of religious orthodoxy in his absolute rejection of group selection. David Sloan Wilson has supplemented Durkheim’s view of religion (as being primarily about group cohesion) with evolutionary analyses to propose that religion was the conduit that pulled humans through a “major transition” in evolutionary history.

Dawkins, along with George Williams and most critics of group selection, acknowledge that natural selection works on groups as well as on individuals, and that group selection is possible in principle. But Dawkins relies on Williams’ argument that selection pressures at the individual level are, in practice, always stronger than those at the group level: free riders will always undercut Darwin’s suggestion that morality evolved because virtuous groups outcompeted selfish groups.

Wilson, however, in *Darwin’s Cathedral*, makes the case that culture in general and religion in particular change the variables in Williams’ analysis. Religions and their associated practices greatly
increase the costs of defection (through punishment and ostracism), increase the contributions of individuals to group efforts (through cultural and emotional mechanisms that increase trust), and sharpen the boundaries — biological and cultural — between groups. Throw in recent discoveries that genetic evolution can work much faster than previously supposed, and the widely respected work of Pete Richerson and Rob Boyd on cultural group selection, and suddenly the old consensus against group selection is outdated.

It’s time to examine the question anew. Yet Dawkins has referred to group selection in interviews as a "heresy," and in *The God Delusion* he dismisses it without giving a reason. In chapter 5 he states the standard Williams free rider objection, notes the argument that religion is a way around the Williams objection, concedes that Darwin believed in group selection, and then moves on. Dismissing a credible position without reasons, and calling it a heresy (even if tongue in cheek), are hallmarks of standard moral thinking, not scientific thinking.

4) Morality is about more than harm and fairness. In *Letter to a Christian Nation*, Sam Harris gives us a standard liberal definition of morality: "Questions of morality are questions about happiness and suffering... To the degree that our actions can affect the experience of other creatures positively or negatively, questions of morality apply." He then goes on to show that the Bible and the Koran, taken literally, are immoral books because they’re not primarily about happiness and suffering, and in many places they advocate harming people.

Reading Harris is like watching professional wrestling or the Harlem Globetrotters. It’s great fun, with lots of acrobatics, but it must not be mistaken for an actual contest. If we want to stage a fair fight between religious and secular moralities, we can’t eliminate one by definition before the match begins. So here’s my definition of morality, which gives each side a chance to make its case:

*Moral systems are interlocking sets of values, practices, institutions, and evolved psychological mechanisms that work together to suppress or regulate selfishness and make social life possible.*

In my research I have found that there are two common ways that cultures suppress and regulate selfishness, two visions of what society is and how it ought to work. I’ll call them the *contractual* approach and the *beehive* approach.

The contractual approach takes the individual as the fundamental unit of value. The fundamental problem of social life is that individuals often hurt each other, and so we create implicit social contracts and explicit laws to foster a fair, free, and safe society in which individuals can pursue their interests and develop themselves and their relationships as they choose.

Morality is about happiness and suffering (as Harris says, and as John Stuart Mill said before him), and so contractualists are endlessly trying to fine-tune laws, reinvent institutions, and extend new rights as circumstances change in order to maximize happiness and minimize suffering. To build a contractual morality, all you need are the two individualizing foundations: harm/care, and fairness/reciprocity. The other three foundations, and any religion that builds on them, run afoul of the prime directive: let people make their own choices, as long as they harm nobody else.

The beehive approach, in contrast, takes the group and its territory as fundamental sources of value. Individual bees are born and die by the thousands, but the hive lives for a long time, and each individual has a role to play in fostering its success. The two fundamental problems of social life are attacks from outside and subversion from within. Either one can lead to the death of the hive, so all
Jonathan Haidt: Moral Psychology and the Misunderstanding of Religion

must pull together, do their duty, and be willing to make sacrifices for the group. Bees don’t have to learn how to behave in this way but human children do, and this is why cultural conservatives are so heavily focused on what happens in schools, families, and the media.

Conservatives generally have a more pessimistic view of human nature than do liberals. They are more likely to believe that if you stand back and give kids space to grow as they please, they’ll grow into shallow, self-centered, undisciplined pleasure seekers. Cultural conservatives work hard to cultivate moral virtues based on the three binding foundations: ingroup/loyalty, authority/respect, and purity/sanctity, as well as on the universally employed foundations of harm/care and fairness/reciprocity. The beehive ideal is not a world of maximum freedom, it is a world of order and tradition in which people are united by a shared moral code that is effectively enforced, which enables people to trust each other to play their interdependent roles. It is a world of very high social capital and low anomie.

It might seem obvious to you that contractual societies are good, modern, creative and free, whereas beehive societies reek of feudalism, fascism, and patriarchy. And, as a secular liberal I agree that contractual societies such as those of Western Europe offer the best hope for living peacefully together in our increasingly diverse modern nations (although it remains to be seen if Europe can solve its current diversity problems).

I just want to make one point, however, that should give contractualists pause: surveys have long shown that religious believers in the United States are happier, healthier, longer-lived, and more generous to charity and to each other than are secular people. Most of these effects have been documented in Europe too. If you believe that morality is about happiness and suffering, then I think you are obligated to take a close look at the way religious people actually live and ask what they are doing right.

Don’t dismiss religion on the basis of a superficial reading of the Bible and the newspaper. Might religious communities offer us insights into human flourishing? Can they teach us lessons that would improve wellbeing even in a primarily contractualist society.

You can’t use the New Atheists as your guide to these lessons. The new atheists conduct biased reviews of the literature and conclude that there is no good evidence on any benefits except the health benefits of religion. Here is Daniel Dennett in Breaking the Spell on whether religion brings out the best in people:

"Perhaps a survey would show that as a group atheists and agnostics are more respectful of the law, more sensitive to the needs of others, or more ethical than religious people. Certainly no reliable survey has yet been done that shows otherwise. It might be that the best that can be said for religion is that it helps some people achieve the level of citizenship and morality typically found in brights. If you find that conjecture offensive, you need to adjust your perspective. (Breaking the Spell, p. 55.)"

I have italicized the two sections that show ordinary moral thinking rather than scientific thinking. The first is Dennett’s claim not just that there is no evidence, but that there is certainly no evidence, when in fact surveys have shown for decades that religious practice is a strong predictor of charitable giving. Arthur Brooks recently analyzed these data (in Who Really Cares) and concluded that the enormous generosity of religious believers is not just recycled to religious charities.
Religious believers give more money than secular folk to secular charities, and to their neighbors. They give more of their time, too, and of their blood. Even if you excuse secular liberals from charity because they vote for government welfare programs, it is awfully hard to explain why secular liberals give so little blood. The bottom line, Brooks concludes, is that all forms of giving go together, and all are greatly increased by religious participation and slightly increased by conservative ideology (after controlling for religiosity).

These data are complex and perhaps they can be spun the other way, but at the moment it appears that Dennett is wrong in his reading of the literature. Atheists may have many other virtues, but on one of the least controversial and most objective measures of moral behavior—giving time, money, and blood to help strangers in need—religious people appear to be morally superior to secular folk.

My conclusion is not that secular liberal societies should be made more religious and conservative in a utilitarian bid to increase happiness, charity, longevity, and social capital. Too many valuable rights would be at risk, too many people would be excluded, and societies are so complex that it's impossible to do such social engineering and get only what you bargained for. My point is just that every longstanding ideology and way of life contains some wisdom, some insights into ways of suppressing selfishness, enhancing cooperation, and ultimately enhancing human flourishing.

But because of the four principles of moral psychology it is extremely difficult for people, even scientists, to find that wisdom once hostilities erupt. A militant form of atheism that claims the backing of science and encourages "brights" to take up arms may perhaps advance atheism. But it may also backfire, polluting the scientific study of religion with moralistic dogma and damaging the prestige of science in the process.

Downloaded from:
http://edge.org/conversation/moral-psychology-and-the-misunderstanding-of-religion