

# The Moral Instinct

Which of the following people would you say is the most admirable: Mother Teresa, Bill Gates or Norman Borlaug? And which do you think is the least admirable? For most people, it's an easy question. Mother Teresa, famous for ministering to the poor in Calcutta, has been beatified by the Vatican, awarded the Nobel Peace Prize and ranked in an American poll as the most admired person of the 20th century. Bill Gates, infamous for giving us the Microsoft dancing paper clip and the blue screen of death, has been decapitated in effigy in "I Hate Gates" Web sites and hit with a pie in the face. As for Norman Borlaug . . . who the heck is Norman Borlaug?



Illustration by Adrian Tomine

Yet a deeper look might lead you to rethink your answers. Borlaug, father of the "Green Revolution" that used agricultural science to reduce world hunger, has been credited with saving a billion lives, more than anyone else in history. Gates, in deciding what to do with his fortune, crunched the numbers and determined that he could alleviate the most misery by fighting everyday scourges in the developing world like malaria, diarrhea and parasites. Mother Teresa, for her part, extolled the virtue of suffering and ran her well-financed missions accordingly: their sick patrons were offered plenty of prayer but harsh conditions, few analgesics and dangerously primitive medical care.

It's not hard to see why the moral reputations of this trio should be so out of line with the good they have done. Mother Teresa was the very embodiment of saintliness: white-clad, sad-eyed, ascetic and often photographed with the wretched of the earth. Gates is a nerd's nerd and the world's richest man, as likely to enter heaven as the proverbial camel squeezing through the needle's eye. And Borlaug, now 93, is an agronomist who has spent his life in labs and nonprofits, seldom walking onto the media stage, and hence into our consciousness, at all.

I doubt these examples will persuade anyone to favor Bill Gates over Mother Teresa for sainthood. But they show that our heads can be turned by an aura of sanctity, distracting us from a more objective reckoning of the actions that make people suffer or flourish. It seems we may all be vulnerable to moral illusions the ethical equivalent of the bending lines that trick the eye on cereal boxes and in psychology textbooks.

Illusions are a favorite tool of perception scientists for exposing the workings of the five senses, and of philosophers for shaking people out of the naïve belief that our minds give us a transparent window onto the world (since if our eyes can be fooled by an illusion, why should we trust them at other times?). Today, a new field is using illusions to unmask a sixth sense, the moral sense. Moral intuitions are being drawn out of people in the lab, on Web sites and in brain scanners, and are being explained with tools from game theory, neuroscience and evolutionary biology.

“Two things fill the mind with ever new and increasing admiration and awe, the oftener and more steadily we reflect on them,” wrote Immanuel Kant, “the starry heavens above and the moral law within.” These days, the moral law within is being viewed with increasing awe, if not always admiration. The human moral sense turns out to be an organ of considerable complexity, with quirks that reflect its evolutionary history and its neurobiological foundations.

These quirks are bound to have implications for the human predicament. Morality is not just any old topic in psychology but close to our conception of the meaning of life. Moral goodness is what gives each of us the sense that we are worthy human beings. We seek it in our friends and mates, nurture it in our children, advance it in our politics and justify it with our religions. A disrespect for morality is blamed for everyday sins and history’s worst atrocities. To carry this weight, the concept of morality would have to be bigger than any of us and outside all of us.

So dissecting moral intuitions is no small matter. If morality is a mere trick of the brain, some may fear, our very grounds for being moral could be eroded. Yet as we shall see, the science of the moral sense can instead be seen as a way to strengthen those grounds, by clarifying what morality is and how it should steer our actions.

### *The Moralization Switch*

The starting point for appreciating that there *is* a distinctive part of our psychology for morality is seeing how moral judgments differ from other kinds of opinions we have on how people ought to behave. Moralization is a psychological state that can be turned on and off like a switch, and when it is on, a distinctive mind-set commandeers our thinking. This is the mind-set that makes us deem actions immoral (“killing is wrong”), rather than merely disagreeable (“I hate brussels sprouts”), unfashionable (“bell-bottoms are out”) or imprudent (“don’t scratch mosquito bites”).

The first hallmark of moralization is that the rules it invokes are felt to be universal. Prohibitions of [rape](#) and murder, for example, are felt not to be matters of local custom but to be universally and objectively warranted. One can easily say, “I don’t like brussels sprouts, but I don’t care if you eat them,” but no one would say, “I don’t like killing, but I don’t care if you murder someone.”

The other hallmark is that people feel that those who commit immoral acts deserve to be punished. Not only is it allowable to inflict pain on a person who has broken a moral rule; it is wrong *not* to, to “let them get away with it.” People are thus untroubled in inviting divine retribution or the power of the state to harm other people they deem immoral. Bertrand Russell wrote, “The infliction of cruelty with a good conscience is a delight to moralists — that is why they invented hell.”

*Steven Pinker is the Johnstone Family Professor of Psychology at Harvard University and the author of “The Language Instinct” and “The Stuff of Thought: Language as a Window Into Human Nature.”*

(Page 2 of 8)

We all know what it feels like when the moralization switch flips inside us — the righteous glow, the burning dudgeon, the drive to recruit others to the cause. The psychologist Paul Rozin has studied the toggle switch by comparing two kinds of people who engage in the same behavior but with different switch settings. Health [vegetarians](#) avoid meat for practical reasons, like lowering [cholesterol](#) and avoiding toxins. Moral vegetarians avoid meat for ethical reasons: to avoid complicity in the suffering of animals. By investigating their feelings about meat-eating, Rozin showed that the moral motive sets off a cascade of opinions. Moral vegetarians are more likely to treat meat as a contaminant — they refuse, for example, to eat a bowl of soup into which a drop of beef broth has fallen. They are more likely to think that other people ought to be vegetarians, and are more likely to imbue their dietary habits with other virtues, like believing that meat avoidance makes people less aggressive and bestial.

Much of our recent social history, including the culture wars between liberals and conservatives, consists of the moralization or amoralization of particular kinds of behavior. Even when people agree that an outcome is desirable, they may disagree on whether it should be treated as a matter of preference and prudence or as a matter of sin and virtue. Rozin notes, for example, that [smoking](#) has lately been moralized. Until recently, it was understood that some people didn’t enjoy smoking or avoided it

because it was hazardous to their health. But with the discovery of the harmful effects of secondhand smoke, smoking is now treated as immoral. Smokers are ostracized; images of people smoking are censored; and entities touched by smoke are felt to be contaminated (so hotels have not only nonsmoking rooms but nonsmoking *floors*). The desire for retribution has been visited on tobacco companies, who have been slapped with staggering “punitive damages.”

At the same time, many behaviors have been amoralized, switched from moral failings to lifestyle choices. They include divorce, illegitimacy, being a working mother, marijuana use and homosexuality. Many afflictions have been reassigned from payback for bad choices to unlucky misfortunes. There used to be people called “bums” and “tramps”; today they are “homeless.” **Drug addiction** is a “disease”; **syphilis** was rebranded from the price of wanton behavior to a “**sexually transmitted disease**” and more recently a “sexually transmitted infection.”

This wave of amoralization has led the cultural right to lament that morality itself is under assault, as we see in the group that anointed itself the Moral Majority. In fact there seems to be a Law of Conservation of Moralization, so that as old behaviors are taken out of the moralized column, new ones are added to it. Dozens of things that past generations treated as practical matters are now ethical battlegrounds, including disposable diapers, I.Q. tests, poultry farms, Barbie dolls and research on **breast cancer**. Food alone has become a minefield, with critics sermonizing about the size of sodas, the chemistry of **fat**, the freedom of chickens, the price of coffee beans, the species of fish and now the distance the food has traveled from farm to plate.

Many of these moralizations, like the assault on smoking, may be understood as practical tactics to reduce some recently identified harm. But whether an activity flips our mental switches to the “moral” setting isn’t just a matter of how much harm it does. We don’t show contempt to the man who fails to change the batteries in his smoke alarms or takes his family on a driving vacation, both of which multiply the risk they will die in an accident. Driving a gas-guzzling Hummer is reprehensible, but driving a gas-guzzling old Volvo is not; eating a Big Mac is unconscionable, but not imported cheese or crême brûlée. The reason for these double standards is obvious: people tend to align their moralization with their own lifestyles.

### *Reasoning and Rationalizing*

It’s not just the content of our moral judgments that is often questionable, but the way

we arrive at them. We like to think that when we have a conviction, there are good reasons that drove us to adopt it. That is why an older approach to moral psychology, led by Jean Piaget and Lawrence Kohlberg, tried to document the lines of reasoning that guided people to moral conclusions. But consider these situations, originally devised by the psychologist Jonathan Haidt:

Julie is traveling in France on summer vacation from college with her brother Mark. One night they decide that it would be interesting and fun if they tried making love. Julie was already taking birth-control pills, but Mark uses a [condom](#), too, just to be safe. They both enjoy the sex but decide not to do it again. They keep the night as a special secret, which makes them feel closer to each other. What do you think about that — was it O.K. for them to make love?

A woman is cleaning out her closet and she finds her old American flag. She doesn't want the flag anymore, so she cuts it up into pieces and uses the rags to clean her bathroom.

A family's dog is killed by a car in front of their house. They heard that dog meat was delicious, so they cut up the dog's body and cook it and eat it for dinner.

Most people immediately declare that these acts are wrong and then grope to justify *why* they are wrong. It's not so easy. In the case of Julie and Mark, people raise the possibility of children with [birth defects](#), but they are reminded that the couple were diligent about [contraception](#). They suggest that the siblings will be emotionally hurt, but the story makes it clear that they weren't. They submit that the act would offend the community, but then recall that it was kept a secret. Eventually many people admit, "I don't know, I can't explain it, I just know it's wrong." People don't generally engage in moral reasoning, Haidt argues, but moral *rationalization*: they begin with the conclusion, coughed up by an unconscious emotion, and then work backward to a plausible justification.

The gap between people's convictions and their justifications is also on display in the favorite new sandbox for moral [psychologists](#), a thought experiment devised by the philosophers Philippa Foot and Judith Jarvis Thomson called the Trolley Problem. On your morning walk, you see a trolley car hurtling down the track, the conductor slumped over the controls. In the path of the trolley are five men working on the track, oblivious to the danger. You are standing at a fork in the track and can pull a lever that will divert the trolley onto a spur, saving the five men. Unfortunately, the trolley

would then run over a single worker who is laboring on the spur. Is it permissible to throw the switch, killing one man to save five? Almost everyone says “yes.”

Consider now a different scene. You are on a bridge overlooking the tracks and have spotted the runaway trolley bearing down on the five workers. Now the only way to stop the trolley is to throw a heavy object in its path. And the only heavy object within reach is a fat man standing next to you. Should you throw the man off the bridge? Both dilemmas present you with the option of sacrificing one life to save five, and so, by the utilitarian standard of what would result in the greatest good for the greatest number, the two dilemmas are morally equivalent. But most people don't see it that way: though they would pull the switch in the first dilemma, they would not heave the fat man in the second. When pressed for a reason, they can't come up with anything coherent, though moral philosophers haven't had an easy time coming up with a relevant difference, either.

*Steven Pinker is the Johnstone Family Professor of Psychology at Harvard University and the author of “The Language Instinct” and “The Stuff of Thought: Language as a Window Into Human Nature.”*

(Page 3 of 8)

When psychologists say “most people” they usually mean “most of the two dozen sophomores who filled out a questionnaire for beer money.” But in this case it means most of the 200,000 people from a hundred countries who shared their intuitions on a Web-based experiment conducted by the psychologists Fiery Cushman and Liane Young and the biologist Marc Hauser. A difference between the acceptability of switch-pulling and man-heaving, and an inability to justify the choice, was found in respondents from Europe, Asia and North and South America; among men and women, blacks and whites, teenagers and octogenarians, Hindus, Muslims, Buddhists, Christians, Jews and atheists; people with elementary-school educations and people with Ph.D.'s.

Joshua Greene, a philosopher and cognitive neuroscientist, suggests that evolution equipped people with a revulsion to manhandling an innocent person. This instinct, he suggests, tends to overwhelm any utilitarian calculus that would tot up the lives saved and lost. The impulse against roughing up a fellow human would explain other examples in which people abjure killing one to save many, like euthanizing a hospital patient to harvest his organs and save five dying patients in need of transplants, or

throwing someone out of a crowded lifeboat to keep it afloat.

By itself this would be no more than a plausible story, but Greene teamed up with the cognitive neuroscientist Jonathan Cohen and several Princeton colleagues to peer into people's brains using functional [M.R.I.](#) They sought to find signs of a conflict between brain areas associated with emotion (the ones that recoil from harming someone) and areas dedicated to rational analysis (the ones that calculate lives lost and saved).

When people pondered the dilemmas that required killing someone with their bare hands, several networks in their brains lighted up. One, which included the medial (inward-facing) parts of the frontal lobes, has been implicated in emotions about other people. A second, the dorsolateral (upper and outer-facing) surface of the frontal lobes, has been implicated in ongoing mental computation (including nonmoral reasoning, like deciding whether to get somewhere by plane or train). And a third region, the anterior cingulate cortex (an evolutionarily ancient strip lying at the base of the inner surface of each cerebral hemisphere), registers a conflict between an urge coming from one part of the brain and an advisory coming from another.

But when the people were pondering a hands-off dilemma, like switching the trolley onto the spur with the single worker, the brain reacted differently: only the area involved in rational calculation stood out. Other studies have shown that neurological patients who have blunted emotions because of damage to the frontal lobes become utilitarians: they think it makes perfect sense to throw the fat man off the bridge. Together, the findings corroborate Greene's theory that our nonutilitarian intuitions come from the victory of an emotional impulse over a cost-benefit analysis.

### *A Universal Morality?*

The findings of trolleyology — complex, instinctive and worldwide moral intuitions — led Hauser and John Mikhail (a legal scholar) to revive an analogy from the philosopher John Rawls between the moral sense and language. According to [Noam Chomsky](#), we are born with a “universal grammar” that forces us to analyze speech in terms of its grammatical structure, with no conscious awareness of the rules in play. By analogy, we are born with a universal moral grammar that forces us to analyze human action in terms of its moral structure, with just as little awareness.

The idea that the moral sense is an innate part of human nature is not far-fetched. A list of human universals collected by the anthropologist Donald E. Brown includes

many moral concepts and emotions, including a distinction between right and wrong; empathy; fairness; admiration of generosity; rights and obligations; proscription of murder, rape and other forms of violence; redress of wrongs; sanctions for wrongs against the community; shame; and taboos.

The stirrings of morality emerge early in childhood. Toddlers spontaneously offer toys and help to others and try to comfort people they see in distress. And according to the psychologists Elliot Turiel and Judith Smetana, preschoolers have an inkling of the difference between societal conventions and moral principles. Four-year-olds say that it is not O.K. to wear pajamas to school (a convention) and also not O.K. to hit a little girl for no reason (a moral principle). But when asked whether these actions would be O.K. if the teacher allowed them, most of the children said that wearing pajamas would now be fine but that hitting a little girl would still not be.

*Steven Pinker is the Johnstone Family Professor of Psychology at Harvard University and the author of "The Language Instinct" and "The Stuff of Thought: Language as a Window Into Human Nature."*

(Page 4 of 8)

Though no one has identified genes for morality, there is circumstantial evidence they exist. The character traits called "conscientiousness" and "agreeableness" are far more correlated in identical **twins** separated at birth (who share their genes but not their environment) than in adoptive siblings raised together (who share their environment but not their genes). People given diagnoses of "**antisocial personality disorder**" or "psychopathy" show signs of morality **blindness** from the time they are children. They bully younger children, torture animals, habitually lie and seem incapable of empathy or remorse, often despite normal family backgrounds. Some of these children grow up into the monsters who bilk elderly people out of their savings, rape a succession of women or shoot convenience-store clerks lying on the floor during a robbery.

Though psychopathy probably comes from a genetic predisposition, a milder version can be caused by damage to frontal regions of the brain (including the areas that inhibit intact people from throwing the hypothetical fat man off the bridge). The neuroscientists Hanna and Antonio Damasio and their colleagues found that some children who sustain severe injuries to their frontal lobes can grow up into callous and irresponsible adults, despite normal intelligence. They lie, steal, ignore **punishment**, endanger their own children and can't think through even the simplest moral

dilemmas, like what two people should do if they disagreed on which TV channel to watch or whether a man ought to steal a drug to save his dying wife.

The moral sense, then, may be rooted in the design of the normal human brain. Yet for all the awe that may fill our minds when we reflect on an innate moral law within, the idea is at best incomplete. Consider this moral dilemma: A runaway trolley is about to kill a schoolteacher. You can divert the trolley onto a sidetrack, but the trolley would trip a switch sending a signal to a class of 6-year-olds, giving them permission to name a teddy bear Muhammad. Is it permissible to pull the lever?

This is no joke. Last month a British woman teaching in a private school in Sudan allowed her class to name a teddy bear after the most popular boy in the class, who bore the name of the founder of Islam. She was jailed for blasphemy and threatened with a public flogging, while a mob outside the prison demanded her death. To the protesters, the woman's life clearly had less value than maximizing the dignity of their religion, and their judgment on whether it is right to divert the hypothetical trolley would have differed from ours. Whatever grammar guides people's moral judgments can't be all *that* universal. Anyone who stayed awake through Anthropology 101 can offer many other examples.

Of course, languages vary, too. In Chomsky's theory, languages conform to an abstract blueprint, like having phrases built out of verbs and objects, while the details vary, like whether the verb or the object comes first. Could we be wired with an abstract spec sheet that embraces all the strange ideas that people in different cultures moralize?

### *The Varieties of Moral Experience*

When anthropologists like Richard Shweder and Alan Fiske survey moral concerns across the globe, they find that a few themes keep popping up from amid the diversity. People everywhere, at least in some circumstances and with certain other folks in mind, think it's bad to harm others and good to help them. They have a sense of fairness: that one should reciprocate favors, reward benefactors and punish cheaters. They value loyalty to a group, sharing and solidarity among its members and conformity to its norms. They believe that it is right to defer to legitimate authorities and to respect people with high status. And they exalt purity, cleanliness and sanctity while loathing defilement, contamination and carnality.

The exact number of themes depends on whether you're a lumper or a splitter, but Haidt counts five — harm, fairness, community (or group loyalty), authority and purity — and suggests that they are the primary colors of our moral sense. Not only do they keep reappearing in cross-cultural surveys, but each one tugs on the moral intuitions of people in our own culture. Haidt asks us to consider how much money someone would have to pay us to do hypothetical acts like the following:

Stick a pin into your palm.

Stick a pin into the palm of a child you don't know. (Harm.)

Accept a wide-screen TV from a friend who received it at no charge because of a computer error.

Accept a wide-screen TV from a friend who received it from a thief who had stolen it from a wealthy family. (Fairness.)

Say something bad about your nation (which you don't believe) on a talk-radio show in your nation.

Say something bad about your nation (which you don't believe) on a talk-radio show in a foreign nation. (Community.)

Slap a friend in the face, with his permission, as part of a comedy skit.

Slap your minister in the face, with his permission, as part of a comedy skit. (Authority.)

Attend a performance-art piece in which the actors act like idiots for 30 minutes, including flubbing simple problems and falling down on stage.

Attend a performance-art piece in which the actors act like animals for 30 minutes, including crawling around naked and urinating on stage. (Purity.)

In each pair, the second action feels far more repugnant. Most of the moral illusions we have visited come from an unwarranted intrusion of one of the moral spheres into our judgments. A violation of community led people to frown on using an old flag to clean a bathroom. Violations of purity repelled the people who judged the morality of consensual incest and prevented the moral vegetarians and nonsmokers from

tolerating the slightest trace of a vile contaminant. At the other end of the scale, displays of extreme purity lead people to venerate religious leaders who dress in white and affect an aura of chastity and asceticism.

### ***The Genealogy of Morals***

*Steven Pinker is the Johnstone Family Professor of Psychology at Harvard University and the author of "The Language Instinct" and "The Stuff of Thought: Language as a Window Into Human Nature."*