The New York Times nytimes.com

January 13, 2008

The Moral Instinct

By STEVEN PINKER

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?

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."

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 crème 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.

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.

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

The five spheres are good candidates for a periodic table of the moral sense not only because they are ubiquitous but also because they appear to have deep evolutionary roots. The impulse to avoid harm, which gives trolley ponderers the willies when they consider throwing a man off a bridge, can also be found in rhesus monkeys, who go hungry rather than pull a chain that delivers food to them and a shock to another monkey. Respect for authority is clearly related to the pecking orders of dominance and appeasement that are widespread in the animal kingdom. The purity-defilement contrast taps the emotion of disgust that is triggered by potential disease

vectors like bodily effluvia, decaying flesh and unconventional forms of meat, and by risky sexual practices like incest.

The other two moralized spheres match up with the classic examples of how altruism can evolve that were worked out by sociobiologists in the 1960s and 1970s and made famous by Richard Dawkins in his book "The Selfish Gene." Fairness is very close to what scientists call reciprocal altruism, where a willingness to be nice to others can evolve as long as the favor helps the recipient more than it costs the giver and the recipient returns the favor when fortunes reverse. The analysis makes it sound as if reciprocal altruism comes out of a robotlike calculation, but in fact Robert Trivers, the biologist who devised the theory, argued that it is implemented in the brain as a suite of moral emotions. Sympathy prompts a person to offer the first favor, particularly to someone in need for whom it would go the furthest. Anger protects a person against cheaters who accept a favor without reciprocating, by impelling him to punish the ingrate or sever the relationship. Gratitude impels a beneficiary to reward those who helped him in the past. Guilt prompts a cheater in danger of being found out to repair the relationship by redressing the misdeed and advertising that he will behave better in the future (consistent with Mencken's definition of *conscience* as "the inner voice which warns us that someone might be looking"). Many experiments on who helps whom, who likes whom, who punishes whom and who feels guilty about what have confirmed these predictions.

Community, the very different emotion that prompts people to share and sacrifice without an expectation of payback, may be rooted in nepotistic altruism, the empathy and solidarity we feel toward our relatives (and which evolved because any gene that pushed an organism to aid a relative would have helped copies of itself sitting inside that relative). In humans, of course, communal feelings can be lavished on nonrelatives as well. Sometimes it pays people (in an evolutionary sense) to love their companions because their interests are yoked, like spouses with common children, in-laws with common relatives, friends with common tastes or allies with common enemies. And sometimes it doesn't pay them at all, but their kinship-detectors have been tricked into treating their groupmates as if they were relatives by tactics like kinship metaphors (*blood brothers*, *fraternities*, *the fatherland*), origin myths, communal meals and other bonding rituals.

Juggling the Spheres

All this brings us to a theory of how the moral sense can be universal and variable at the same time. The five moral spheres are universal, a legacy of evolution. But how they are ranked in importance, and which is brought in to moralize which area of social life — sex, government, commerce, religion, diet and so on — depends on the culture. Many of the flabbergasting practices in faraway places become more intelligible when you recognize that the same

moralizing impulse that Western elites channel toward violations of harm and fairness (our moral obsessions) is channeled elsewhere to violations in the other spheres. Think of the Japanese fear of nonconformity (community), the holy ablutions and dietary restrictions of Hindus and Orthodox Jews (purity), the outrage at insulting the Prophet among Muslims (authority). In the West, we believe that in business and government, fairness should trump community and try to root out nepotism and cronyism. In other parts of the world this is incomprehensible — what heartless creep would favor a perfect stranger over his own brother?

The ranking and placement of moral spheres also divides the cultures of liberals and conservatives in the United States. Many bones of contention, like homosexuality, atheism and one-parent families from the right, or racial imbalances, sweatshops and executive pay from the left, reflect different weightings of the spheres. In a large Web survey, Haidt found that liberals put a lopsided moral weight on harm and fairness while playing down group loyalty, authority and purity. Conservatives instead place a moderately high weight on all five. It's not surprising that each side thinks it is driven by lofty ethical values and that the other side is base and unprincipled.

Reassigning an activity to a different sphere, or taking it out of the moral spheres altogether, isn't easy. People think that a behavior belongs in its sphere as a matter of sacred necessity and that the very act of questioning an assignment is a moral outrage. The psychologist Philip Tetlock has shown that the mentality of taboo — a conviction that some thoughts are sinful to think — is not just a superstition of Polynesians but a mind-set that can easily be triggered in college-educated Americans. Just ask them to think about applying the sphere of reciprocity to relationships customarily governed by community or authority. When Tetlock asked subjects for their opinions on whether adoption agencies should place children with the couples willing to pay the most, whether people should have the right to sell their organs and whether they should be able to buy their way out of jury duty, the subjects not only disagreed but felt personally insulted and were outraged that anyone would raise the question.

The institutions of modernity often question and experiment with the way activities are assigned to moral spheres. Market economies tend to put everything up for sale. Science amoralizes the world by seeking to understand phenomena rather than pass judgment on them. Secular philosophy is in the business of scrutinizing all beliefs, including those entrenched by authority and tradition. It's not surprising that these institutions are often seen to be morally corrosive.

Is Nothing Sacred?

And "morally corrosive" is exactly the term that some critics would apply to the new science of

the moral sense. The attempt to dissect our moral intuitions can look like an attempt to debunk them. Evolutionary psychologists seem to want to unmask our noblest motives as ultimately self-interested — to show that our love for children, compassion for the unfortunate and sense of justice are just tactics in a Darwinian struggle to perpetuate our genes. The explanation of how different cultures appeal to different spheres could lead to a spineless relativism, in which we would never have grounds to criticize the practice of another culture, no matter how barbaric, because "we have our kind of morality and they have theirs." And the whole enterprise seems to be dragging us to an amoral nihilism, in which morality itself would be demoted from a transcendent principle to a figment of our neural circuitry.

In reality, none of these fears are warranted, and it's important to see why not. The first misunderstanding involves the logic of evolutionary explanations. Evolutionary biologists sometimes anthropomorphize DNA for the same reason that science teachers find it useful to have their students imagine the world from the viewpoint of a molecule or a beam of light. One shortcut to understanding the theory of selection without working through the math is to imagine that the genes are little agents that try to make copies of themselves.

Unfortunately, the meme of the selfish gene escaped from popular biology books and mutated into the idea that organisms (including people) are ruthlessly self-serving. And this doesn't follow. Genes are not a reservoir of our dark unconscious wishes. "Selfish" genes are perfectly compatible with selfless organisms, because a gene's metaphorical goal of selfishly replicating itself can be implemented by wiring up the brain of the organism to do unselfish things, like being nice to relatives or doing good deeds for needy strangers. When a mother stays up all night comforting a sick child, the genes that endowed her with that tenderness were "selfish" in a metaphorical sense, but by no stretch of the imagination is *she* being selfish.

Nor does reciprocal altruism — the evolutionary rationale behind fairness — imply that people do good deeds in the cynical expectation of repayment down the line. We all know of unrequited good deeds, like tipping a waitress in a city you will never visit again and falling on a grenade to save platoonmates. These bursts of goodness are not as anomalous to a biologist as they might appear.

In his classic 1971 article, Trivers, the biologist, showed how natural selection could push in the direction of true selflessness. The emergence of tit-for-tat reciprocity, which lets organisms trade favors without being cheated, is just a first step. A favor-giver not only has to avoid blatant cheaters (those who would accept a favor but not return it) but also prefer generous reciprocators (those who return the biggest favor they can afford) over stingy ones (those who return the smallest favor they can get away with). Since it's good to be chosen as a recipient of favors, a competition arises to be the most generous partner around. More accurately, a

competition arises to *appear* to be the most generous partner around, since the favor-giver can't literally read minds or see into the future. A reputation for fairness and generosity becomes an asset.

Now this just sets up a competition for potential beneficiaries to inflate their reputations without making the sacrifices to back them up. But it also pressures the favor-giver to develop ever-more-sensitive radar to distinguish the genuinely generous partners from the hypocrites. This arms race will eventually reach a logical conclusion. The most effective way to *seem* generous and fair, under harsh scrutiny, is to be generous and fair. In the long run, then, reputation can be secured only by commitment. At least some agents evolve to be genuinely high-minded and self-sacrificing — they are moral not because of what it brings them but because that's the kind of people they are.

Of course, a theory that predicted that everyone always sacrificed themselves for another's good would be as preposterous as a theory that predicted that no one ever did. Alongside the niches for saints there are niches for more grudging reciprocators, who attract fewer and poorer partners but don't make the sacrifices necessary for a sterling reputation. And both may coexist with outright cheaters, who exploit the unwary in one-shot encounters. An ecosystem of niches, each with a distinct strategy, can evolve when the payoff of each strategy depends on how many players are playing the other strategies. The human social environment does have its share of generous, grudging and crooked characters, and the genetic variation in personality seems to bear the fingerprints of this evolutionary process.

Is Morality a Figment?

So a biological understanding of the moral sense does not entail that people are calculating maximizers of their genes or self-interest. But where does it leave the concept of morality itself?

Here is the worry. The scientific outlook has taught us that some parts of our subjective experience are products of our biological makeup and have no objective counterpart in the world. The qualitative difference between red and green, the tastiness of fruit and foulness of carrion, the scariness of heights and prettiness of flowers are design features of our common nervous system, and if our species had evolved in a different ecosystem or if we were missing a few genes, our reactions could go the other way. Now, if the distinction between right and wrong is also a product of brain wiring, why should we believe it is any more real than the distinction between red and green? And if it is just a collective hallucination, how could we argue that evils like genocide and slavery are wrong for everyone, rather than just distasteful to us?

Putting God in charge of morality is one way to solve the problem, of course, but Plato made

short work of it 2,400 years ago. Does God have a good reason for designating certain acts as moral and others as immoral? If not — if his dictates are divine whims — why should we take them seriously? Suppose that God commanded us to torture a child. Would that make it all right, or would some other standard give us reasons to resist? And if, on the other hand, God was forced by moral reasons to issue some dictates and not others — if a command to torture a child was never an option — then why not appeal to those reasons directly?

This throws us back to wondering where those reasons could come from, if they are more than just figments of our brains. They certainly aren't in the physical world like wavelength or mass. The only other option is that moral truths exist in some abstract Platonic realm, there for us to discover, perhaps in the same way that mathematical truths (according to most mathematicians) are there for us to discover. On this analogy, we are born with a rudimentary concept of number, but as soon as we build on it with formal mathematical reasoning, the nature of mathematical reality forces us to discover some truths and not others. (No one who understands the concept of two, the concept of four and the concept of addition can come to any conclusion but that 2 + 2 = 4.) Perhaps we are born with a rudimentary moral sense, and as soon as we build on it with moral reasoning, the nature of moral reality forces us to some conclusions but not others.

Moral realism, as this idea is called, is too rich for many philosophers' blood. Yet a diluted version of the idea — if not a list of cosmically inscribed Thou-Shalts, then at least a few If-Thens — is not crazy. Two features of reality point any rational, self-preserving social agent in a moral direction. And they could provide a benchmark for determining when the judgments of our moral sense are aligned with morality itself.

One is the prevalence of nonzero-sum games. In many arenas of life, two parties are objectively better off if they both act in a nonselfish way than if each of them acts selfishly. You and I are both better off if we share our surpluses, rescue each other's children in danger and refrain from shooting at each other, compared with hoarding our surpluses while they rot, letting the other's child drown while we file our nails or feuding like the Hatfields and McCoys. Granted, I might be a bit better off if I acted selfishly at your expense and you played the sucker, but the same is true for you with me, so if each of us tried for these advantages, we'd both end up worse off. Any neutral observer, and you and I if we could talk it over rationally, would have to conclude that the state we should aim for is the one in which we both are unselfish. These spreadsheet projections are not quirks of brain wiring, nor are they dictated by a supernatural power; they are in the nature of things.

The other external support for morality is a feature of rationality itself: that it cannot depend on the egocentric vantage point of the reasoner. If I appeal to you to do anything that affects me

— to get off my foot, or tell me the time or not run me over with your car — then I can't do it in a way that privileges my interests over yours (say, retaining my right to run you over with my car) if I want you to take me seriously. Unless I am Galactic Overlord, I have to state my case in a way that would force me to treat you in kind. I can't act as if my interests are special just because I'm me and you're not, any more than I can persuade you that the spot I am standing on is a special place in the universe just because I happen to be standing on it.

Not coincidentally, the core of this idea — the interchangeability of perspectives — keeps reappearing in history's best-thought-through moral philosophies, including the Golden Rule (itself discovered many times); Spinoza's Viewpoint of Eternity; the Social Contract of Hobbes, Rousseau and Locke; Kant's Categorical Imperative; and Rawls's Veil of Ignorance. It also underlies Peter Singer's theory of the Expanding Circle — the optimistic proposal that our moral sense, though shaped by evolution to overvalue self, kin and clan, can propel us on a path of moral progress, as our reasoning forces us to generalize it to larger and larger circles of sentient beings.

Doing Better by Knowing Ourselves

Morality, then, is still something larger than our inherited moral sense, and the new science of the moral sense does not make moral reasoning and conviction obsolete. At the same time, its implications for our moral universe are profound.

At the very least, the science tells us that even when our adversaries' agenda is most baffling, they may not be amoral psychopaths but in the throes of a moral mind-set that appears to them to be every bit as mandatory and universal as ours does to us. Of course, some adversaries really are psychopaths, and others are so poisoned by a punitive moralization that they are beyond the pale of reason. (The actor Will Smith had many historians on his side when he recently speculated to the press that Hitler thought he was acting morally.) But in any conflict in which a meeting of the minds is not completely hopeless, a recognition that the other guy is acting from moral rather than venal reasons can be a first patch of common ground. One side can acknowledge the other's concern for community or stability or fairness or dignity, even while arguing that some other value should trump it in that instance. With affirmative action, for example, the opponents can be seen as arguing from a sense of fairness, not racism, and the defenders can be seen as acting from a concern with community, not bureaucratic power. Liberals can ratify conservatives' concern with families while noting that gay marriage is perfectly consistent with that concern.

The science of the moral sense also alerts us to ways in which our psychological makeup can get in the way of our arriving at the most defensible moral conclusions. The moral sense, we are

learning, is as vulnerable to illusions as the other senses. It is apt to confuse morality per se with purity, status and conformity. It tends to reframe practical problems as moral crusades and thus see their solution in punitive aggression. It imposes taboos that make certain ideas indiscussible. And it has the nasty habit of always putting the self on the side of the angels.

Though wise people have long reflected on how we can be blinded by our own sanctimony, our public discourse still fails to discount it appropriately. In the worst cases, the thoughtlessness of our brute intuitions can be celebrated as a virtue. In his influential essay "The Wisdom of Repugnance," Leon Kass, former chair of the President's Council on Bioethics, argued that we should disregard reason when it comes to cloning and other biomedical technologies and go with our gut: "We are repelled by the prospect of cloning human beings . . . because we intuit and feel, immediately and without argument, the violation of things that we rightfully hold dear. . . . In this age in which everything is held to be permissible so long as it is freely done . . . repugnance may be the only voice left that speaks up to defend the central core of our humanity. Shallow are the souls that have forgotten how to shudder."

There are, of course, good reasons to regulate human cloning, but the shudder test is not one of them. People have shuddered at all kinds of morally irrelevant violations of purity in their culture: touching an untouchable, drinking from the same water fountain as a Negro, allowing Jewish blood to mix with Aryan blood, tolerating sodomy between consenting men. And if our ancestors' repugnance had carried the day, we never would have had autopsies, vaccinations, blood transfusions, artificial insemination, organ transplants and in vitro fertilization, all of which were denounced as immoral when they were new.

There are many other issues for which we are too quick to hit the moralization button and look for villains rather than bug fixes. What should we do when a hospital patient is killed by a nurse who administers the wrong drug in a patient's intravenous line? Should we make it easier to sue the hospital for damages? Or should we redesign the IV fittings so that it's physically impossible to connect the wrong bottle to the line?

And nowhere is moralization more of a hazard than in our greatest global challenge. The threat of human-induced climate change has become the occasion for a moralistic revival meeting. In many discussions, the cause of climate change is overindulgence (too many S.U.V.'s) and defilement (sullying the atmosphere), and the solution is temperance (conservation) and expiation (buying carbon offset coupons). Yet the experts agree that these numbers don't add up: even if every last American became conscientious about his or her carbon emissions, the effects on climate change would be trifling, if for no other reason than that two billion Indians and Chinese are unlikely to copy our born-again abstemiousness. Though voluntary conservation may be one wedge in an effective carbon-reduction pie, the other wedges will have

to be morally boring, like a carbon tax and new energy technologies, or even taboo, like nuclear power and deliberate manipulation of the ocean and atmosphere. Our habit of moralizing problems, merging them with intuitions of purity and contamination, and resting content when we feel the right feelings, can get in the way of doing the right thing.

Far from debunking morality, then, the science of the moral sense can advance it, by allowing us to see through the illusions that evolution and culture have saddled us with and to focus on goals we can share and defend. As Anton Chekhov wrote, "Man will become better when you show him what he is like."

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."

Copyright 2008 The New York Times Company

Privacy Policy | Search | Corrections | RSS | First Look | Help | Contact Us | Work for Us | Site Map