

Discover<sup>1</sup>

April 21 2004

## *Whose Life Would You Save?*

*Scientists say morality may be hardwired into our brains by evolution*

by Carl Zimmer,

Illustrations by Jeff Westby

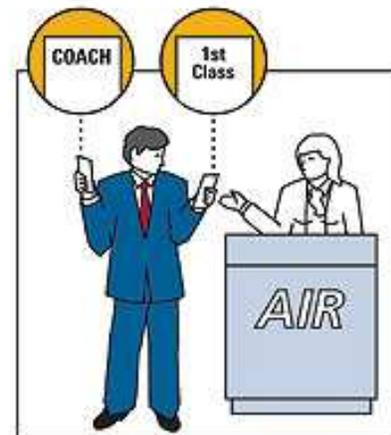
Dinner with a philosopher is never just dinner, even when it's at an obscure Indian restaurant on a quiet side street in Princeton with a 30-year-old postdoctoral researcher. Joshua Greene is a man who spends his days thinking about right and wrong and how we separate the two. He has a particular fondness for moral paradoxes, which he collects the way some people collect snow globes.

"Let's say you're walking by a pond and there's a drowning baby," Greene says, over chicken tikka masala. "If you said, 'I've just paid \$200 for these shoes and the water would ruin them, so I won't save the baby,' you'd be an awful, horrible person. But there are millions of children around the world in the same situation, where just a little money for medicine or food could save their lives. And yet we don't consider ourselves monsters for having this dinner rather than giving the money to Oxfam. Why is that?"

Philosophers pose this sort of puzzle over dinner every day. What's unusual here is what Greene does next to sort out the conundrum. He leaves the restaurant, walks down Nassau Street to the building that houses Princeton University's psychology department, and says hello to graduate student volunteer Nishant Patel. (Greene's volunteers take part in his study anonymously; Patel is not his real name.) They walk downstairs to the basement, where Patel dumps his keys and wallet and shoes in a basket. Greene waves an airport metal-detector paddle up and down Patel's legs, then guides him into an adjoining room dominated by a magnetic resonance imaging scanner. The student lies down on a slab, and Greene closes a cagelike device over his head. Pressing a button, Greene maneuvers Patel's head into a massive doughnut-shaped magnet.

Greene goes back to the control room to calibrate the MRI, then begins to send Patel messages. They are beamed into the scanner by a video projector and bounce off a mirror just above Patel's nose. Among the messages that Greene sends is the following dilemma, cribbed from the final episode of the TV series *M\*A\*S\*H*: A group of villagers is hiding in a basement while enemy soldiers search the rooms above. Suddenly, a baby among them starts to cry. The villagers know that if the soldiers hear it they will come in and kill everyone. "Is it appropriate," the message reads, "for you to smother your child in order to save yourself and the other villagers?"

As Patel ponders this question—and others like it—the MRI scans his brain, revealing crackling clusters of neurons. Over the past four years, Greene has scanned dozens of people making these kinds of moral judgments. What he has found can be unsettling. Most of us would like to believe that when we say something is right or wrong, we are using our powers of reason alone. But Greene argues that our emotions also play a powerful role in our moral judgments, triggering instinctive responses that are the product of



You are checking in for a flight when the person at the counter accidentally gives you a boarding pass for a first-class seat. Your ticket is for coach.

Do you point out the mistake?

"Is it appropriate," the message reads, "for you to smother your child in order to save yourself and the other villagers?"

millions of years of evolution. “A lot of our deeply felt moral convictions may be quirks of our evolutionary history,” he says.

Greene’s research has put him at the leading edge of a field so young it still lacks an official name. Moral neuroscience? Neuroethics? Whatever you call it, the promise is profound. “Some people in these experiments think we’re putting their soul under the microscope,” Greene says, “and in a sense, that is what we’re doing.”

The puzzle of moral judgments grabbed Greene’s attention when he was a philosophy major at Harvard University. Most modern theories of moral reasoning, he learned, were powerfully shaped by one of two great philosophers: Immanuel Kant and John Stuart Mill. Kant believed that pure reason alone could lead us to moral truths. Based on his own pure reasoning, for instance, he declared that it was wrong to use someone for your own ends and that it was right to act only according to principles that everyone could follow.

John Stuart Mill, by contrast, argued that the rules of right and wrong should above all else achieve the greatest good for the greatest number of people, even though particular individuals might be worse off as a result. (This approach became known as utilitarianism, based on the “utility” of a moral rule.) “Kant puts what’s right before what’s good,” says Greene. “Mill puts what’s good before what’s right.”

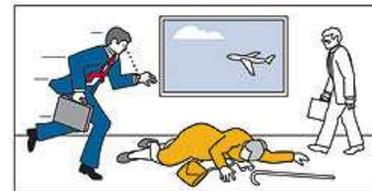
By the time Greene came to Princeton for graduate school in 1997, however, he had become dissatisfied with utilitarians and Kantians alike. None of them could explain how moral judgments work in the real world. Consider, for example, this thought experiment concocted by the philosophers Judith Jarvis Thompson and Philippa Foot: Imagine you’re at the wheel of a trolley and the brakes have failed. You’re approaching a fork in the track at top speed. On the left side, five rail workers are fixing the track. On the right side, there is a single worker. If you do nothing, the trolley will bear left and kill the five workers. The only way to save five lives is to take the responsibility for changing the trolley’s path by hitting a switch. Then you will kill one worker. What would you do?

Now imagine that you are watching the runaway trolley from a footbridge. This time there is no fork in the track. Instead, five workers are on it, facing certain death. But you happen to be standing next to a big man. If you sneak up on him and push him off the footbridge, he will fall to his death. Because he is so big, he will stop the trolley. Do you willfully kill one man, or do you allow five people to die?

Logically, the questions have similar answers. Yet if you poll your friends, you’ll probably find that many more are willing to throw a switch than push someone off a bridge. It is hard to explain why what seems right in one case can seem wrong in another. Sometimes we act more like Kant and sometimes more like Mill. “The trolley problem seemed to boil that conflict down to its essence,” Greene says. “If I could figure out how to make sense of that particular problem, I could make sense of the whole Kant-versus-Mill problem in ethics.”

The crux of the matter, Greene decided, lay not in the logic of moral judgments but in the role our emotions play in forming them. He began to explore the psychological studies of the 18th-century Scottish philosopher David Hume. Hume argued that people call an act good not because they rationally determine it to be so but because it makes them feel good. They call an act bad because it fills them with disgust. Moral knowledge, Hume wrote, comes partly from an “immediate feeling and finer internal sense.”

Moral instincts have deep roots, primatologists have found. Last September, for instance, Sarah Brosnan and Frans de Waal of Emory University reported that monkeys have a sense of fairness. Brosnan and De Waal trained capuchin monkeys to take a pebble from them; if the monkeys gave the pebble back, they got a cucumber. Then they ran the same experiment with two monkeys sitting in adjacent cages so that each could see the other. One monkey still got a cucumber, but the other one got a grape—a tastier reward. More



You are running down a crowded corridor in the airport, trying to catch a flight that’s about to leave. Suddenly, an old woman in front of you slips and falls hard.

Do you stop to help, knowing that you’ll miss your plane?

than half the monkeys who got cucumbers balked at the exchange. Sometimes they threw the cucumber at the researchers; sometimes they refused to give the pebble back. Apparently, De Waal says, they realized that they weren't being treated fairly.

In an earlier study, De Waal observed a colony of chimpanzees that got fed by their zookeeper only after they had all gathered in an enclosure. One day, a few young chimps dallied outside for hours, leaving the rest to go hungry. The next day, the other chimps attacked the stragglers, apparently to punish them for their selfishness. The primates seemed capable of moral judgment without benefit of human reasoning. "Chimps may be smart," Greene says. "But they don't read Kant."

The evolutionary origins of morality are easy to imagine in a social species. A sense of fairness would have helped early primates cooperate. A sense of disgust and anger at cheaters would have helped them avoid falling into squabbling. As our ancestors became more self-aware and acquired language, they would transform those feelings into moral codes that they then taught their children.

This idea made a lot of sense to Greene. For one thing, it showed how moral judgments can feel so real. "We make moral judgments so automatically that we don't really understand how they're formed," he says. It also offered a potential solution to the trolley problem: Although the two scenarios have similar outcomes, they trigger different circuits in the brain. Killing someone with your bare hands would most likely have been recognized as immoral millions of years ago. It summons ancient and overwhelmingly negative emotions—despite any good that may come of the killing. It simply feels wrong.

Throwing a switch for a trolley, on the other hand, is not the sort of thing our ancestors confronted. Cause and effect, in this case, are separated by a chain of machines and electrons, so they do not trigger a snap moral judgment. Instead, we rely more on abstract reasoning—weighing costs and benefits, for example—to choose between right and wrong. Or so Greene hypothesized. When he arrived at Princeton, he had no way to look inside people's brains. Then in 1999, Greene learned that the university was building a brain-imaging center.

The heart of the Center for the Study of Brain, Mind, and Behavior is an MRI scanner in the basement of Green Hall. The scanner creates images of the brain by generating an intense magnetic field. Some of the molecules in the brain line up with the field, and the scanner wiggles the field back and forth a few degrees. As the molecules wiggle, they release radio waves. By detecting the waves, the scanner can reconstruct the brain as well as detect where neurons are consuming oxygen—a sign of mental activity. In two seconds, the center's scanner can pinpoint such activity down to a cubic millimeter—about the size of a peppercorn.

When neuroscientists first started scanning brains in the early 1990s, they studied the basic building blocks of thought, such as language, vision, and attention. But in recent years, they've also tried to understand how the brain works when people interact. Humans turn out to have special neural networks that give them what many cognitive neuroscientists call social intelligence. Some regions can respond to smiles, frowns, and other expressions in a tenth of a second. Others help us get inside a person's head and figure out intentions. When neuroscientist Jonathan Cohen came to Princeton to head the center, he hoped he could dedicate some time with the scanner to study the interaction between cognition and emotion. Greene's morality study was a perfect fit.

Working with Cohen and other scientists at the center, Greene decided to compare how the brain responds to different questions. He took the trolley problem as his starting point, then invented questions



Your plane has been taken over by a terrorist. He has taken a passenger's baby hostage and is holding a knife to his throat.

Do you rush the terrorist to subdue him, knowing that the baby will die before you get there?

designed to place volunteers on a spectrum of moral judgment. Some questions involved personal moral choices; some were impersonal but no less moral. Others were utterly innocuous, such as deciding whether to take a train or a bus to work. Greene could then peel away the brain's general decision-making circuits and focus in on the neural patterns that differentiate personal from impersonal thought.

Some scenarios were awful, but Greene suspected people would make quick decisions about them. Should you kill a friend's sick father so he can collect on the insurance policy? Of course not. But other questions—like the one about the smothered baby—were as agonizing as they were gruesome. Greene calls these doozies. "If they weren't creepy, we wouldn't be doing our job," he says.

As Greene's subjects mulled over his questions, the scanner measured the activity in their brains. When all the questions had flashed before the volunteers, Greene was left with gigabytes of data, which then had to be mapped onto a picture of the brain. "It's not hard, like philosophy hard, but there are so many details to keep track of," he says. When he was done, he experienced a "pitter-patter heartbeat moment." Just as he had predicted, personal moral decisions tended to stimulate certain parts of the brain more than impersonal moral decisions.

The more people Greene scanned, the clearer the pattern became: Impersonal moral decisions (like whether to throw a switch on a trolley) triggered many of the same parts of the brain as nonmoral questions do (such as whether you should take the train or the bus to work). Among the regions that became active was a patch on the surface of the brain near the temples. This region, known as the dorsolateral prefrontal cortex, is vital for logical thinking. Neuroscientists believe it helps keep track of several pieces of information at once so that they can be compared. "We're using our brains to make decisions about things that evolution hasn't wired us up for," Greene says.

Personal moral questions lit up other areas. One, located in the cleft of the brain behind the center of the forehead, plays a crucial role in understanding what other people are thinking or feeling. A second, known as the superior temporal sulcus, is located just above the ear; it gathers information about people from the way they move their lips, eyes, and hands. A third, made up of parts of two adjacent regions known as the posterior cingulate and the precuneus, becomes active when people feel strong emotions.

Greene suspects these regions are part of a neural network that produces the emotional instincts behind many of our moral judgments. The superior temporal sulcus may help make us aware of others who would be harmed. Mind reading lets us appreciate their suffering. The precuneus may help trigger a negative feeling—an inarticulate sense, for example, that killing someone is plain wrong.

When Greene and his coworkers first began their study, not a single scan of the brain's moral decision-making process had been published. Now a number of other scientists are investigating the neural basis of morality, and their results are converging on some of the same ideas. "The neuroanatomy seems to be coming together," Greene says.

Another team of neuroscientists at Princeton, for instance, has pinpointed neural circuits that govern the sense of fairness. Economists have known for a long time that humans, like capuchin monkeys, get annoyed to an irrational degree when they feel they're getting shortchanged. A classic example of this phenomenon crops up during the "ultimatum game," in which two players are given a chance to split some money. One player proposes the split, and the other can accept or reject it—but if he rejects it, neither player gets anything.

If both players act in a purely rational way, as most economists assume people act, the game should have a predictable result. The first player will offer the second the worst possible split, and the second will be obliged to accept it. A little money, after all, is better than none. But in experiment after experiment, players tend to offer something close to a 50-50 split. Even more remarkably, when they offer significantly less than half, they're often rejected.

The Princeton team (led by Alan Sanfey, now at the University of Arizona) sought to explain that rejection by having people play the ultimatum game while in the MRI scanner. Their subjects always played the part of the responder. In some cases the proposer was another person; in others it was a computer. Sanfey

found that unfair offers from human players—more than those from the computer—triggered pronounced reactions in a strip of the brain called the anterior insula.

Previous studies had shown that this area produces feelings of anger and disgust. The stronger the response, Sanfey and his colleagues found, the more likely the subject would reject the offer.

Another way to study moral intuition is to look at brains that lack it. James Blair at the National Institute of Mental Health has spent years performing psychological tests on criminal psychopaths. He has found that they have some puzzling gaps in perception. They can put themselves inside the heads of other people, for example, acknowledging that others feel fear or sadness. But they have a hard time recognizing fear or sadness, either on people's faces or in their voices.

Blair says that the roots of criminal psychopathy can first be seen in childhood. An abnormal level of neurotransmitters might make children less empathetic. When most children see others get sad or angry, it disturbs them and makes them want to avoid acting in ways that provoke such reactions. But budding psychopaths don't perceive other people's pain, so they don't learn to rein in their violent outbursts.

Your plane has made an emergency landing on the ocean. All but one lifeboat has been destroyed, and it's so full it's beginning to sink. Should you throw people overboard, beginning with the elderly who are too weak to resist, to save yourself and the others?
--

As Greene's database grows, he can see more clearly how the brain's intuitive and reasoning networks are activated. In most cases, one dominates the other. Sometimes, though, they produce opposite responses of equal strength, and the brain has difficulty choosing between them. Part of the evidence for this lies in the time it takes for Greene's volunteers to answer his questions. Impersonal moral ones and nonmoral ones tend to take about the same time to answer. But when people decide that personally hurting or killing someone is appropriate, it takes them a long time to say yes—twice as long as saying no to these particular kinds of questions. The brain's emotional network says no, Greene's brain scans show, and its reasoning network says yes.

When two areas of the brain come into conflict, researchers have found, an area known as the anterior cingulate cortex, or ACC, switches on to mediate between them. Psychologists can trigger the ACC with a simple game called the Stroop test, in which people have to name the color of a word. If subjects are shown the word blue in red letters, for instance, their responses slow down and the ACC lights up. "It's the area of the brain that says, 'Hey, we've got a problem here,'" Greene says.

Greene's questions, it turns out, pose a sort of moral Stroop test. In cases where people take a long time to answer agonizing personal moral questions, the ACC becomes active. "We predicted that we'd see this, and that's what we got," he says. Greene, in other words, may be exposing the biology of moral anguish.

Of course, not all people feel the same sort of moral anguish. Nor do they all answer Greene's questions the same way. Some aren't willing to push a man over a bridge, but others are. Greene nicknames these two types the Kantians and the utilitarians. As he takes more scans, he hopes to find patterns of brain activity that are unique to each group. "This is what I've wanted to get at from the beginning," Greene says, "to understand what makes some people do some things and other people do other things."

Greene knows that his results can be disturbing: "People sometimes say to me, 'If everyone believed what you say, the whole world would fall apart.'" ' If right and wrong are nothing more than the instinctive firing of neurons, why bother being good? But Greene insists the evidence coming from neuroimaging can't be ignored. "Once you understand someone's behavior on a sufficiently mechanical level, it's very hard to look at them as evil," he says. "You can look at them as dangerous; you can pity them. But evil doesn't exist on a neuronal level."

By the time Patel emerges from the scanner, rubbing his eyes, it's past 11 p.m. "I can try to print a copy of your brain now or e-mail it to you later," Greene says. Patel looks at the image on the computer screen and decides to pass. "This doesn't feel like you?" Greene says with a sly smile. "You're not going to send this to your mom?"

Soon Greene and Patel, who is Indian, are talking about whether Indians and Americans might answer some moral questions differently. All human societies share certain moral universals, such as fairness and sympathy. But Greene argues that different cultures produce different kinds of moral intuition and different kinds of brains. Indian morality, for instance, focuses more on matters of purity, whereas American morality focuses on individual autonomy. Researchers such as Jonathan Haidt, a psychologist at the University of Virginia, suggest that such differences shape a child's brain at a relatively early age. By the time we become adults, we're wired with emotional responses that guide our judgments for the rest of our lives.

Many of the world's great conflicts may be rooted in such neuronal differences, Greene says, which may explain why the conflicts seem so intractable. "We have people who are talking past each other, thinking the other people are either incredibly dumb or willfully blind to what's right in front of them," Greene says. "It's not just that people disagree; it's that they have a hard time imagining how anyone could disagree on this point that seems so obvious." Some people wonder how anyone could possibly tolerate abortion. Others wonder how women could possibly go out in public without covering their faces. The answer may be that their brains simply don't work the same: Genes, culture, and personal experience have wired their moral circuitry in different patterns.

Greene hopes that research on the brain's moral circuitry may ultimately help resolve some of these seemingly irresolvable disputes. "When you have this understanding, you have a bit of distance between yourself and your gut reaction," he says. "You may not abandon your core values, but it makes you a more reasonable person. Instead of saying, 'I am right, and you are just nuts,' you say, 'This is what I care about, and we have a conflict of interest we have to work around.'"

Greene could go on — that's what philosophers do — but he needs to switch back to being a neuroscientist. It's already late, and Patel's brain will take hours to decode.

## Notes

<sup>1</sup><http://discovermagazine.com/2004/apr/whose-life-would-you-save>