

The Roots of Morality

Neurobiologists, philosophers, psychologists, and legal scholars are probing the nature of human morality using a variety of experimental techniques and moral challenges

A team of psychologists recently asked dozens of college students to consider several morally charged situations. In one, a friend lies on his résumé to land a job; in another, survivors of a plane crash consider cannibalizing an injured boy to avoid starvation. Students who pondered these hypothetical scenarios while sitting at a filthy desk with sticky stains and a chewed-up pen rated them as more immoral than did students who sat at a pristine desk. In another version of the experiment, a nearby trash can doused with novelty fart spray had a similar effect. The findings, in press at *Personality and Social Psychology Bulletin*, demonstrate that emotions such as disgust exert a powerful influence on moral judgments, even when they are triggered by something unrelated to the moral issue, says study co-author Jonathan Haidt, a psychologist at the University of Virginia, Charlottesville.

Haidt is one of a growing number of researchers taking an experimental approach to investigating the nature of

human morality. The field has drawn practitioners from diverse backgrounds including philosophy, psychology, and neuroscience. They don't always see eye to eye, but they are united in their belief that the scientific method will yield fresh insights into questions that have vexed philosophers for centuries.

One area of intense interest is the interplay of emotion and reason in moral decision-making. Haidt argues that people rely on gut reactions to tell right from wrong and employ reason mainly when they try to justify their intuitions after the fact, not unlike an art museum visitor who is struck by the beauty of a painting but struggles to explain why. Not everyone accepts this view, but other researchers do see evidence that moral judgments are surprisingly automatic. "I think there is an emerging consensus that things happen pretty quickly and that explicit conscious reasoning is not where the action is," Haidt says.

This automaticity has led some researchers to suggest that the human brain has built-in

moral instincts. Cognitive neuroscientists are already hunting for the underlying neural mechanisms. At the same time, psychologists and anthropologists are searching for evidence of universal moral principles shared by all people. Others are interested in how morality differs from culture to culture. They are using techniques that include brain imaging and online questionnaires to probe the roots of morality, and some researchers are viewing the development of moral principles through the lens of evolution.

The work is likely to yield a better understanding of our moral intuitions and where they come from, says Walter Sinnott-Armstrong, a philosopher at Dartmouth College. Philosophers, from the ancient Greeks on, have tried to answer these questions mainly through introspection, an exercise that has often amounted to seeking new arguments for a previously held conviction, says Sinnott-Armstrong, who has recently begun some experimental work of his own. "One thing that's fascinating about science is you don't know where you're going to end up."

CREDIT: PETER HOEY

Dissecting moral cognition

Two 18th century thinkers have had a huge influence on moral philosophy: David Hume, a Scotsman, who argued that passions drive moral judgments, and Immanuel Kant, a German, who countered that dispassionate reason is, or ought to be, the driving force. The clash between these two philosophical titans still reverberates today.

Lately, Hume seems to be gaining an edge, thanks to the work of Haidt and others. In an influential 2001 paper in *Psychological Review*, Haidt describes an experiment in which he and colleagues asked people to consider a hypothetical situation involving a brother and sister who decide to have sex. They use two forms of birth control, enjoy the experiment, but decide not to do it again. Most people took little time to condemn the siblings' actions as morally wrong but then struggled when pressed to explain why. After all, there was virtually no chance of conception and the vignette had made it clear that the siblings were not emotionally scarred by the experience. Many of the volunteers eventually resorted to an explanation along the lines of "I just know it's wrong." If people were reasoning their way to an opinion, Haidt argued, they wouldn't be so dumbfounded when asked to explain it.

In more recent work, Haidt has investigated whether manipulating emotions can alter moral judgments. The messy desk experiment suggests that it can, as does an earlier study in which Haidt and then-graduate student Thalia Wheatley used hypnotic suggestion to trigger a wave of disgust in volunteers as they read vignettes about morally dubious behavior. Volunteers issued harsher moral judgments for vignettes containing a cue word that triggered the hypnotic suggestion than they did for an alternative version with slightly different wording. Wheatley and Haidt reported in 2005 in *Psychological Science*.

Disgust even raised people's moral suspicions when the act described was innocuous. One scenario described a student council member picking topics for faculty-student discussions. When this vignette contained the disgust-triggering cue word, subjects rated the student's activities as less morally appropriate. "It just seems like he's up to something," one wrote.

Other evidence that emotions guide moral judgments comes from work with people who've suffered damage to

brain regions that mediate emotion. In a 2007 paper in *Nature*, a team led by Michael Koenigs of the University of Iowa, Iowa City, and Antonio Damasio of the University of Southern California in Los Angeles reported that people with damage to the ventromedial prefrontal cortex made abnormal judgments on hypothetical moral dilemmas that forced them to consider whether it was permissible to sacrifice the life of one person to save several others. These scenarios included variants of the so-called trolley problem, a favorite tool of morality researchers. One version puts the subject behind the wheel of a runaway trolley headed toward five hapless workers; the only way to save the five is to hit a switch on the dashboard that would divert the trolley to a track with just one worker. Healthy volunteers and lesion patients alike tended to say this was acceptable. The two groups differed, however, on a more emotionally charged version of the dilemma in which the only way to save the five is to shove a large man off a footbridge to stop the runaway trolley. Although the same utilitarian logic applies—kill one to save five—healthy subjects found this option harder to stomach: only about 20% said it would be permissible. But twice as many of the brain-damaged subjects said they would shove the man, suggesting that their damaged emotional circuitry made them unusually likely to pick the utilitarian option.

Jorge Moll, a neuroscientist at Labs D'Or Hospital Network, a private medical and research institute in Rio de Janeiro, Brazil, views the ventromedial prefrontal cortex as part of a network of brain regions underlying "prosocial sentiments" such as guilt and compassion. Moll and colleagues reported last year in *Social Neuroscience* that this brain region is activated by viewing morally evocative photographs, such as ones of a hungry child, even when no judgment is required. In a

2006 paper in the *Proceedings of the National Academy of Sciences (PNAS)*, he and others reported that the same region is activated when volunteers elect to donate money to charity. Moll views prosocial sentiments as the core of morality and thinks they arose from ancient mechanisms that evolved to enable our ancestors to form social attachments and cooperative groups.

The Koenigs study contains hints that emotions aren't the entire story, however, says co-author Marc Hauser, a cognitive scientist at Harvard University. He points out that the lesion patients still made normal judgments in many situations, particularly regarding dilemmas that didn't tug at the emotions and "easier" ones that are emotionally charged but elicit strong consensus among healthy subjects—that it's wrong, for example, to earn money to feed your family by allowing your young daughter to appear in a pornographic film, even in hard times. "That rules out the strong version of the hypothesis that emotions are causally necessary for making [all] moral judgments," Hauser says. "That just can't be right."

Don't get all emotional

An alternative view, championed by Joshua Greene, a cognitive neuroscientist and philosopher at Harvard, is that when people grapple with moral dilemmas like the trolley problems, emotion and rationality duke it out in the brain. In Greene's view, the key difference between flipping the switch and shoving the man off the footbridge is that the latter evokes a negative emotional reaction that overrides cold utilitarian logic.

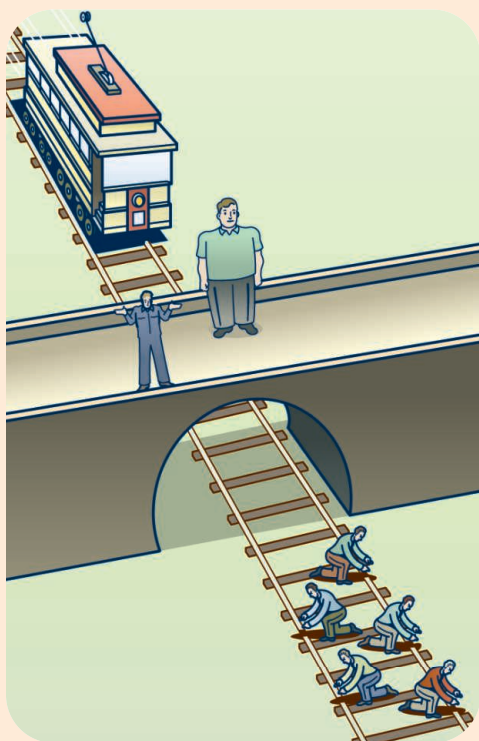
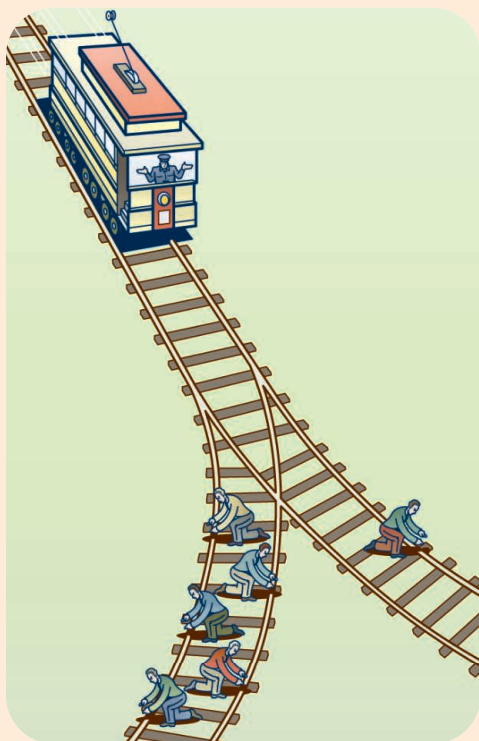
In a 2001 *Science* paper, Greene, then a postdoc with Jonathan Cohen at Princeton University, and colleagues reported that the medial frontal gyrus and other brain regions linked to emotion become more active when people contemplate "personal" moral dilemmas—such as shoving the man onto the trolley tracks or removing a man's

organs against his will to save five transplant recipients—compared with when they weigh impersonal moral dilemmas—such as flipping a switch to save the workers or declaring bogus business expenses on a tax return. These impersonal dilemmas preferentially activate a different set of brain regions thought to contribute to abstract reasoning and problem solving. Greene and colleagues reported in a follow-up study, published in 2004 in *Neuron*.

Based on these findings,



Philosophical difference. New studies tend to support the view of David Hume (left) that emotions drive moral judgments; Immanuel Kant (right) argued that reason should be the driving force.



Moral dilemma. Is it morally acceptable to redirect a runaway trolley car hurtling toward five workers onto a track with just one worker? How about pushing a man off a footbridge into the path of the trolley to stop it before it hits the hapless workers? Most people say they would sacrifice one life to save five in the first scenario but not the second. In this case, emotion may trump utilitarian logic.

Greene envisions a tug of war between emotion and cognition in the brain: Emotions tell us we'll feel terrible if we push the man; cognition says: Push him! Five is greater than one. Greene suspects that the arbiter in this conflict may be a brain region called the anterior cingulate cortex. Previous studies have found that this region fires up when people wrestle with many types of internal conflicts, and it did so when subjects in Greene's study faced particularly difficult moral dilemmas.

In a recent study that mirrors Haidt's work with manipulating emotion, Greene and colleagues had college students evaluate moral dilemmas while grappling with an extra cognitive burden: searching for a particular number in a string of characters scrolling across a computer screen. The extra cognitive work slowed response times when students made utilitarian judgments but not emotional ones, the researchers report in an upcoming issue of *Cognition*. Greene sees the study as evidence that cognition is an important part of moral decision-making.

Getting off track?

Some researchers see the trolley problems as too artificial. "We don't have a lot of faith in using these esoteric examples," says Jordan Grafman, a cognitive neuroscientist at the National Institute of Mental Health in Bethesda, Maryland. The situations are so far-fetched that Grafman and others question whether they really engage the neural mechanisms involved in everyday moral reasoning. Everyday moral reasoning is likely to involve a memory component that's missing in Greene's account, Grafman says. "More often than not, we take a situation we've experienced in the past and compare it to the new one," he says. Brain-imaging studies done with more realistic scenarios might catch some of the underlying neural mechanisms, says Grafman, who is gearing up to do such an experiment in collaboration with Ralph Adolphs and colleagues at the California Institute of Technology in Pasadena, who have been collecting hundreds of real-life moral dilemmas experienced by people of different ages, education levels, and socioeconomic backgrounds. In a

paper published online by *Science* this week (www.sciencemag.org/cgi/content/abstract/1153651), researchers led by Ming Hsu, now at the University of Illinois, Urbana-Champaign, and colleagues at Caltech report taking a different approach: scanning the brains of volunteers as they tried to decide the fairest way to distribute donations to a real-life Ugandan orphanage.

At the same time, some researchers argue that the emphasis on emotion and reason is too simplistic, akin to placing the ghost of Hume in one network of brain regions and the ghost of Kant in another. "It's like they take 18th century categories and try to do 21st century science," says John Mikhail, a legal scholar at Georgetown University in Washington, D.C. Mikhail, Hauser, and others point out that before emotion and reason can evaluate a given situation, the brain has to first answer questions such as who did what to whom, whether someone got hurt, and whether the harm was intentional.

For example, most people would condemn someone who tried to poison a friend's coffee but accidentally stirred in sugar instead of poison. It's the bad intention that matters, not the outcome. To investigate how the brain makes such distinctions, Hauser and Harvard graduate students Liane Young and Fiery Cushman recently teamed up with Rebecca Saxe, a cognitive neuroscientist at the Massachusetts Institute of Technology (MIT) in Cambridge. When volunteers read vignettes about intentional and unintentional harms, activity increased in the right temporoparietal junction (RTPJ), a brain region involved in sussing out other people's intentions. RTPJ activity was greatest for cases like the bungled poisoning in which someone tried but failed to inflict harm, the researchers reported last year in *PNAS*.

At last month's meeting of the Cognitive Neuroscience Society, Saxe and Young reported that interfering with RTPJ activity using a noninvasive method called transcranial magnetic stimulation caused people to downplay intentions and, for example, judge the attempted poisoning less harshly because ultimately no harm was done. Such findings demonstrate that the cognitive contributions to moral judgments aren't limited to the weighing of harms that's emphasized by trolley problems, Saxe says. Understanding intentions is another crucial component, and the RTPJ findings begin to hint at the neural mechanisms involved, she says.

A moral grammar

Some morality researchers see parallels in the study of language, particularly the

influential work of MIT linguist Noam Chomsky, who has argued that humans have an innate capacity for language and that all languages share common principles—a universal grammar. Could there be an analogous moral capacity in the human brain and a universal moral grammar?

Mikhail began pondering these questions as a philosophy graduate student, during a year he spent working with Chomsky at MIT. To investigate, he administered trolley problems and other moral dilemmas to different groups of people, including children and people from non-Western cultures. If there is universal moral grammar, he reasoned, factors such as gender, age, education level, and cultural background should have little influence on the judgments people make. Preliminary results pointed in that direction, and Mikhail's initial work has been expanded and confirmed by Hauser, Cushman, and Young, who developed an online Moral Sense Test (moral.wjh.harvard.edu) that has been taken by more than 200,000 people from 120 countries. Chinese, Spanish, and Dutch versions are now up and running as well, and Hauser is collaborating with several anthropologists to gather similar data from remote indigenous populations in Guatemala, Papua New Guinea, Tanzania, and Bolivia. It's work in progress, Hauser says, but so far "it's looking like there's a lot of similarity across widely different cultures."

Mikhail, meanwhile, has been studying legal texts for clues to what the elements of a universal moral grammar might be. "The law is the one institution in most societies that's responsible for the practical matter of solving day-to-day moral problems that arise," Mikhail says. "The rules of law that have evolved over time, to my mind, are a really good first approximation of the unconscious rules that people use in moral judgments."

Flavors of morality

Although harm and fairness have been the focus of most research so far on the psychology and neuroscience of morality, some researchers think there's more to the story. Haidt argues for five psychological foundations of morality: He includes harm and fairness and adds loyalty, respect for authority, and spiritual purity (*Science*, 18 May 2007, p. 998). Other scholars have proposed lists

of universal aspects of morality, and Haidt identified his five by trying to work out what they all had in common. He hypothesizes that all five exist in every culture but are emphasized to varying degrees. "I see them as being much like the five kinds of taste buds," he says. "If you go around the world, the cuisines differ in how much they rely on each one."

Haidt set up a Web survey (www.YourMorals.org) to evaluate how people weight the five foundations. More than

ferent types of acts morally wrong are based on the same psychological or neurobiological mechanisms." He has been working on brain-imaging experiments to investigate whether different types of moral scenarios engage different neural circuitry.

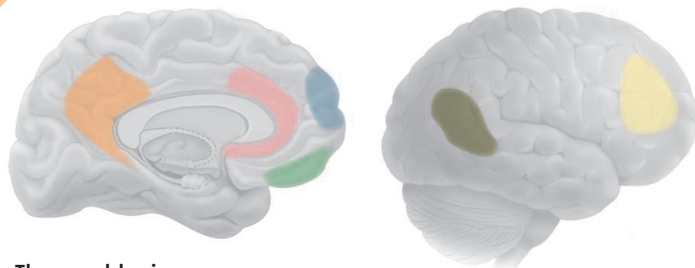
Many researchers think moral cognition depends on neural mechanisms that also play roles in other types of social cognition and are likely present to some degree in our primate kin. Primatologists have found hints of a sense of harm and fairness even in monkeys, who

will forgo food for days to prevent a neighbor from receiving a shock and will reject a small reward when they've learned that a given task usually earns them a larger one. Haidt speculates that morality is an elaboration of primate social behavior that evolved in part because it helped promote cohesiveness in groups of early humans, giving them an advantage over competing groups. Hauser agrees that morality probably has roots in primate social behavior. But that raises a puzzle about why moral decisions seem to feel somehow different, he says. "One of the problems for our field right now is when you say

something is moral, how does the brain know it's moral as opposed to just social?"

It's too early to know where all of the empirical work on morality will lead. Forced to speculate, researchers can envision brain scans that could determine whether a defendant in a murder case had the mental capacity to tell right from wrong and lawyers who wear perfume formulated to sway the emotions—and verdict—of a jury. Sinnott-Armstrong says he can envision revised sentencing guidelines that take human psychology into account. "If we have a better understanding of morality, we'll have a better understanding of how [lawmakers] get their intuitions about how much punishment is deserved," he says. "We might find that [moral intuitions] are more reliable in some cases than in others."

Most likely of all, perhaps, the work may give us a better understanding of ourselves. However, reducing a noble human attribute such as morality to a matter of natural selection and brain activity may lead us into uncomfortable territory, says Saxe: "Even though we know this in hundreds of ways, it continues to be both fascinating and unsettling to find out that something you thought of as a feature of the self turns out to be a product of your brain." **—GREG MILLER**



The moral brain.

Neuroimaging studies have linked several brain regions to moral cognition. Disruptions to the right temporoparietal junction (brown), which is involved in understanding intentions, or the ventromedial prefrontal cortex (green), which processes emotion, have been found to alter moral judgments. Greene and colleagues have suggested that activity in the anterior cingulate cortex (pink) signals conflict between emotion, reflected by activity in the medial frontal gyrus (blue) and other areas (orange, brown), and "cold" cognition, reflected by activity in dorsolateral prefrontal cortex (yellow).

35,000 people have logged on so far, he says, and the findings suggest cultural differences in how people carve up the moral domain. In more liberal cultures, such as Western Europe and Australia, people emphasize harm and fairness over the others. In more conservative cultures, including South Asia and the Middle East, all five foundations are important. In the United States, which falls in the middle of the spectrum, Haidt and colleagues have found a similar divide between self-described liberals and conservatives. Liberals tend to downplay purity, for example, arguing that something can be indecent without being morally wrong, Haidt says. But it's a matter of degree: Although many liberals wonder why conservatives are so hung up on what types of sexual behavior are right and wrong, they have analogous hang-ups, often more symbolic than rational, about food that was processed in certain ways, or by people seen as either villains or victims. Haidt says he hopes the work will spur his colleagues, most of them two-foundation liberals like himself, to think beyond harm and fairness.

Sinnott-Armstrong agrees that morality is multifaceted: "It's not clear to me at all that all those judgments where we call dif-