The Righteous Mind

Why Good People are Divided by Politics and Religion

Jonathan Haidt

Contents

Introduction

"Can we all get along?" That appeal was made famous on May 1, 1992, by Rodney King, a black man who had been beaten nearly to death by four Los Angeles police officers a year earlier. The entire nation had seen a videotape of the beating, so when a jury failed to convict the officers, their acquittal triggered widespread outrage and six days of rioting in Los Angeles. Fifty-three people were killed and more than seven thousand buildings were torched. Much of the mayhem was carried live; news cameras tracked the action from helicopters circling overhead. After a particularly horrific act of violence against a white truck driver, King was moved to make his appeal for peace.

King's appeal is now so overused that it has become cultural kitsch, a catchphrase[1] more often said for laughs than as a serious plea for mutual understanding. I therefore hesitated to use King's words as the opening line of this book, but I decided to go ahead, for two reasons. The first is because most Americans nowadays are asking King's question not about race relations but about political relations and the collapse of cooperation across party lines. Many Americans feel as though the nightly news from Washington is being sent to us from helicopters circling over the city, delivering dispatches from the war zone.

The second reason I decided to open this book with an overused phrase is because King followed it up with something lovely, something rarely quoted. As he stumbled through his television interview, fighting back tears and often repeating himself, he found these words: "Please, we can get along here. We all can get along. I mean, we're all stuck here for a while. Let's try to work it out."

This book is about why it's so hard for us to get along. We are

indeed all stuck here for a while, so let's at least do what we can to understand why we are so easily divided into hostile groups, each one certain of its righteousness.

People who devote their lives to studying something often come to believe that the object of their fascination is the key to understanding everything. Books have been published in recent years on the transformative role in human history played by cooking, mothering, war . . . even salt. This is one of those books. I study moral psychology, and I'm going to make the case that morality is the extraordinary human capacity that made civilization possible. I don't mean to imply that cooking, mothering, war, and salt were not also necessary, but in this book I'm going to take you on a tour of human nature and history from the perspective of moral psychology.

By the end of the tour, I hope to have given you a new way to think about two of the most important, vexing, and divisive topics in human life: politics and religion. Etiquette books tell us not to discuss these topics in polite company, but I say go ahead. Politics and religion are both expressions of our underlying moral psychology, and an understanding of that psychology can help to bring people together. My goal in this book is to drain some of the heat, anger, and divisiveness out of these topics and replace them with awe, wonder, and curiosity. We are downright lucky that we evolved this complex moral psychology that allowed our species to burst out of the forests and savannas and into the delights, comforts, and extraordinary peacefulness of modern societies in just a few thousand years.[2] My hope is that this book will make conversations about morality, politics, and religion more common, more civil, and more fun, even in mixed company. My hope is that it will help us to get along.

BORN TO BE RIGHTEOUS

I could have titled this book *The Moral Mind* to convey the sense that the human mind is designed to "do" morality, just as it's designed to

do language, sexuality, music, and many other things described in popular books reporting the latest scientific findings. But I chose the title *The Righteous Mind* to convey the sense that human nature is not just intrinsically moral, it's also intrinsically moralistic, critical, and judgmental.

The word *righteous* comes from the old Norse word *rettviss* and the old English word *rihtwis*, both of which mean "just, upright, virtuous."[3] This meaning has been carried into the modern English words *righteous* and *righteousness*, although nowadays those words have strong religious connotations because they are usually used to translate the Hebrew word *tzedek*. *Tzedek* is a common word in the Hebrew Bible, often used to describe people who act in accordance with God's wishes, but it is also an attribute of God and of God's judgment of people (which is often harsh but always thought to be just).

The linkage of righteousness and judgmentalism is captured in some modern definitions of *righteous*, such as "arising from an outraged sense of justice, morality, or fair play."[4] The link also appears in the term *self-righteous*, which means "convinced of one's own righteousness, especially in contrast with the actions and beliefs of others; narrowly moralistic and intolerant."[5] I want to show you that an obsession with righteousness (leading inevitably to self-righteousness) is the normal human condition. It is a feature of our evolutionary design, not a bug or error that crept into minds that would otherwise be objective and rational.[6]

Our righteous minds made it possible for human beings—but no other animals—to produce large cooperative groups, tribes, and nations without the glue of kinship. But at the same time, our righteous minds guarantee that our cooperative groups will always be cursed by moralistic strife. Some degree of conflict among groups may even be necessary for the health and development of any society. When I was a teenager I wished for world peace, but now I yearn for a world in which competing ideologies are kept in balance, systems of accountability keep us all from getting away with too much, and fewer people believe that righteous ends justify violent means. Not a very romantic wish, but one that we might actually achieve.

WHAT LIES AHEAD

This book has three parts, which you can think of as three separate books—except that each one depends on the one before it. Each part presents one major principle of moral psychology.

Part I is about the first principle: *Intuitions come first, strategic reasoning second.*[7] Moral intuitions arise automatically and almost instantaneously, long before moral reasoning has a chance to get started, and those first intuitions tend to drive our later reasoning. If you think that moral reasoning is something we do to figure out the truth, you'll be constantly frustrated by how foolish, biased, and illogical people become when they disagree with you. But if you think about moral reasoning as a skill we humans evolved to further our social agendas—to justify our own actions and to defend the teams we belong to—then things will make a lot more sense. Keep your eye on the intuitions, and don't take people's moral arguments at face value. They're mostly post hoc constructions made up on the fly, crafted to advance one or more strategic objectives.

The central metaphor of these four chapters is that *the mind is divided, like a rider on an elephant, and the rider's job is to serve the elephant.* The rider is our conscious reasoning—the stream of words and images of which we are fully aware. The elephant is the other 99 percent of mental processes—the ones that occur outside of awareness but that actually govern most of our behavior.[8] I developed this metaphor in my last book, *The Happiness Hypothesis*, where I described how the rider and elephant work together, sometimes poorly, as we stumble through life in search of meaning and connection. In this book I'll use the metaphor to solve puzzles such as why it seems like everyone (else) is a hypocrite[9] and why political partisans are so willing to believe outrageous lies and conspiracy theories. I'll also use the metaphor to show you how you can better persuade people who seem unresponsive to reason.

Part II is about the second principle of moral psychology, which is that *there's more to morality than harm and fairness.* The central metaphor of these four chapters is that *the righteous mind is like a tongue*

with six taste receptors. Secular Western moralities are like cuisines that try to activate just one or two of these receptors—either concerns about harm and suffering, or concerns about fairness and injustice. But people have so many other powerful moral intuitions, such as those related to liberty, loyalty, authority, and sanctity. I'll explain where these six taste receptors come from, how they form the basis of the world's many moral cuisines, and why politicians on the right have a built-in advantage when it comes to cooking meals that voters like.

Part III is about the third principle: *Morality binds and blinds*. The central metaphor of these four chapters is that *human beings are 90 percent chimp and 10 percent bee*. Human nature was produced by natural selection working at two levels simultaneously. Individuals compete with individuals within every group, and we are the descendants of primates who excelled at that competition. This gives us the ugly side of our nature, the one that is usually featured in books about our evolutionary origins. We are indeed selfish hypocrites so skilled at putting on a show of virtue that we fool even ourselves.

But human nature was also shaped as groups competed with other groups. As Darwin said long ago, the most cohesive and cooperative groups generally beat the groups of selfish individualists. Darwin's ideas about group selection fell out of favor in the 1960s, but recent discoveries are putting his ideas back into play, and the implications are profound. We're not always selfish hypocrites. We also have the ability, under special circumstances, to shut down our petty selves and become like cells in a larger body, or like bees in a hive, working for the good of the group. These experiences are often among the most cherished of our lives, although our hivishness can blind us to other moral concerns. Our bee-like nature facilitates altruism, heroism, war, and genocide.

Once you see our righteous minds as primate minds with a hivish overlay, you get a whole new perspective on morality, politics, and religion. I'll show that our "higher nature" allows us to be profoundly altruistic, but that altruism is mostly aimed at members of our own groups. I'll show that religion is (probably) an evolutionary adaptation for binding groups together and helping them to create

communities with a shared morality. It is not a virus or a parasite, as some scientists (the "New Atheists") have argued in recent years. And I'll use this perspective to explain why some people are conservative, others are liberal (or progressive), and still others become libertarians. People bind themselves into political teams that share moral narratives. Once they accept a particular narrative, they become blind to alternative moral worlds.

(A note on terminology: In the United States, the word *liberal* refers to progressive or left-wing politics, and I will use the word in this sense. But in Europe and elsewhere, the word *liberal* is truer to its original meaning—valuing liberty above all else, including in economic activities. When Europeans use the word *liberal,* they often mean something more like the American term *libertarian,* which cannot be placed easily on the left-right spectrum.[10] Readers from outside the United States may want to swap in the words *progressive* or *left-wing* whenever I say *liberal.*)

In the coming chapters I'll draw on the latest research in neuroscience, genetics, social psychology, and evolutionary modeling, but the take-home message of the book is ancient. It is the realization that we are all self-righteous hypocrites:

> Why do you see the speck in your neighbor's eye, but do not notice the log in your own eye? . . . You hypocrite, first take the log out of your own eye, and then you will see clearly to take the speck out of your neighbor's eye. (MATTHEW 7:3–5)

Enlightenment (or wisdom, if you prefer) requires us all to take the logs out of our own eyes and then escape from our ceaseless, petty, and divisive moralism. As the eighth-century Chinese Zen master Sen-ts'an wrote:

> The Perfect Way is only difficult
> for those who pick and choose;
> Do not like, do not dislike;

all will then be clear.
Make a hairbreadth difference,
and Heaven and Earth are set apart;
If you want the truth to stand clear before you,
never be for or against.
The struggle between "for" and "against"
is the mind's worst disease.[11]

I'm not saying we should live our lives like Sen-ts'an. In fact, I believe that a world without moralism, gossip, and judgment would quickly decay into chaos. But if we want to *understand* ourselves, our divisions, our limits, and our potentials, we need to step back, drop the moralism, apply some moral psychology, and analyze the game we're all playing.

Let us now examine the psychology of this struggle between "for" and "against." It is a struggle that plays out in each of our righteous minds, and among all of our righteous groups.

Intuitions Come First, Strategic Reasoning Second

Central Metaphor

The mind is divided, like a rider on an elephant, and the rider's job is to serve the elephant.

Where Does Morality Come From?

I'm going to tell you a brief story. Pause after you read it and decide whether the people in the story did anything morally wrong.

> A family's dog was killed by a car in front of their house. They had heard that dog meat was delicious, so they cut up the dog's body and cooked it and ate it for dinner. Nobody saw them do this.

If you are like most of the well-educated people in my studies, you felt an initial flash of disgust, but you hesitated before saying the family had done anything *morally* wrong. After all, the dog was dead already, so they didn't hurt it, right? And it was their dog, so they had a right to do what they wanted with the carcass, no? If I pushed you to make a judgment, odds are you'd give me a nuanced answer, something like "Well, I think it's disgusting, and I think they should have just buried the dog, but I wouldn't say it was *morally* wrong."

OK, here's a more challenging story:

> A man goes to the supermarket once a week and buys a chicken. But before cooking the chicken, he

has sexual intercourse with it. Then he cooks it and eats it.

Once again, no harm, nobody else knows, and, like the dog-eating family, it involves a kind of recycling that is—as some of my research subjects pointed out—an efficient use of natural resources. But now the disgust is so much stronger, and the action just seems so . . . degrading. Does that make it wrong? If you're an educated and politically liberal Westerner, you'll probably give another nuanced answer, one that acknowledges the man's right to do what he wants, as long as he doesn't hurt anyone.

But if you are *not* a liberal or libertarian Westerner, you probably think it's wrong—morally wrong—for someone to have sex with a chicken carcass and then eat it. For you, as for most people on the planet, morality is broad. Some actions are wrong even though they don't hurt anyone. Understanding the simple fact that morality differs around the world, and even within societies, is the first step toward understanding your righteous mind. The next step is to understand where these many moralities came from in the first place.

THE ORIGIN OF MORALITY (TAKE 1)

I studied philosophy in college, hoping to figure out the meaning of life. After watching too many Woody Allen movies, I had the mistaken impression that philosophy would be of some help.[1] But I had taken some psychology courses too, and I loved them, so I chose to continue. In 1987 I was admitted to the graduate program in psychology at the University of Pennsylvania. I had a vague plan to conduct experiments on the psychology of humor. I thought it might be fun to do research that let me hang out in comedy clubs.

A week after arriving in Philadelphia, I sat down to talk with Jonathan Baron, a professor who studies how people think and make decisions. With my (minimal) background in philosophy, we had a good discussion about ethics. Baron asked me point-blank: "Is *moral* thinking any different from other kinds of thinking?" I said

that thinking about moral issues (such as whether abortion is wrong) seemed different from thinking about other kinds of questions (such as where to go to dinner tonight), because of the much greater need to provide reasons justifying your moral judgments to other people. Baron responded enthusiastically, and we talked about some ways one might compare moral thinking to other kinds of thinking in the lab. The next day, on the basis of little more than a feeling of encouragement, I asked him to be my advisor and I set off to study moral psychology.

In 1987, moral psychology was a part of developmental psychology. Researchers focused on questions such as how children develop in their thinking about rules, especially rules of fairness. The big question behind this research was: How do children come to know right from wrong? Where does morality come from?

There are two obvious answers to this question: nature or nurture. If you pick nature, then you're a *nativist*. You believe that moral knowledge is native in our minds. It comes preloaded, perhaps in our God-inscribed hearts (as the Bible says), or in our evolved moral emotions (as Darwin argued).[2]

But if you believe that moral knowledge comes from nurture, then you are an *empiricist*.[3] You believe that children are more or less blank slates at birth (as John Locke said).[4] If morality varies around the world and across the centuries, then how could it be innate? Whatever morals we have as adults must have been learned during childhood from our own experience, which includes adults telling us what's right and wrong. (*Empirical* means "from observation or experience.")

But this is a false choice, and in 1987 moral psychology was mostly focused on a third answer: *rationalism*, which says that kids figure out morality for themselves. Jean Piaget, the greatest developmental psychologist of all time, began his career as a zoologist studying mollusks and insects in his native Switzerland. He was fascinated by the stages that animals went through as they transformed themselves from, say, caterpillars to butterflies. Later, when his attention turned to children, he brought with him this interest in stages of development. Piaget wanted to know how the extraordinary sophistication of adult

thinking (a cognitive butterfly) emerges from the limited abilities of young children (lowly caterpillars).

Piaget focused on the kinds of errors kids make. For example, he'd put water into two identical drinking glasses and ask kids to tell him if the glasses held the same amount of water. (Yes.) Then he'd pour the contents of one of the glasses into a tall skinny glass and ask the child to compare the new glass to the one that had not been touched. Kids younger than six or seven usually say that the tall skinny glass now holds more water, because the level is higher. They don't understand that the total volume of water is conserved when it moves from glass to glass. He also found that it's pointless for adults to explain the conservation of volume to kids. The kids won't get it until they reach an age (and cognitive stage) when their minds are ready for it. And when they are ready, they'll figure it out for themselves just by playing with cups of water.

In other words, the understanding of the conservation of volume wasn't innate, and it wasn't learned from adults. Kids *figure it out for themselves*, but only when their minds are ready *and* they are given the right kinds of experiences.

Piaget applied this cognitive-developmental approach to the study of children's moral thinking as well.[5] He got down on his hands and knees to play marbles with children, and sometimes he deliberately broke rules and played dumb. The children then responded to his mistakes, and in so doing, they revealed their growing ability to respect rules, change rules, take turns, and resolve disputes. This growing knowledge came in orderly stages, as children's cognitive abilities matured.

Piaget argued that children's understanding of morality is like their understanding of those water glasses: we can't say that it is innate, and we can't say that kids learn it directly from adults.[6] It is, rather, *self-constructed* as kids play with other kids. Taking turns in a game is like pouring water back and forth between glasses. No matter how often you do it with three-year-olds, they're just not ready to get the concept of fairness,[7] any more than they can understand the conservation of volume. But once they've reached the age of five or six, then playing games, having arguments, and working things out

together will help them learn about fairness far more effectively than any sermon from adults.

This is the essence of psychological rationalism: We grow into our rationality as caterpillars grow into butterflies. If the caterpillar eats enough leaves, it will (eventually) grow wings. And if the child gets enough experiences of turn taking, sharing, and playground justice, it will (eventually) become a moral creature, able to use its rational capacities to solve ever harder problems. Rationality is our nature, and good moral reasoning is the end point of development.

Rationalism has a long and complex history in philosophy. In this book I'll use the word *rationalist* to describe anyone who believes that reasoning is the most important and reliable way to obtain moral knowledge.[8]

Piaget's insights were extended by Lawrence Kohlberg, who revolutionized the study of morality in the 1960s with two key innovations.[9] First, he developed a way to quantify Piaget's observation that children's moral reasoning changed over time. He created a set of moral dilemmas that he presented to children of various ages, and he recorded and coded their responses. For example, should a man named Heinz break into a drugstore to steal a drug that would save his dying wife? Should a girl named Louise reveal to her mother that her younger sister had lied to the mother? It didn't much matter whether the child said yes or no; what mattered were the *reasons* children gave when they tried to explain their answers.

Kohlberg found a six-stage progression in children's reasoning about the *social* world, and this progression matched up well with the stages Piaget had found in children's reasoning about the *physical* world. Young children judged right and wrong by very superficial features, such as whether a person was punished for an action. (If an adult punished the act, then the act must have been wrong.) Kohlberg called the first two stages the "pre-conventional" level of moral judgment, and they correspond to the Piagetian stage at which kids judge the physical world by superficial features (if a glass is taller, then it has more water in it).

But during elementary school, most children move on to the two "conventional" stages, becoming adept at understanding and even

manipulating rules and social conventions. This is the age of petty legalism that most of us who grew up with siblings remember well ("I'm not hitting you. I'm using your hand to hit you. Stop hitting yourself!"). Kids at this stage generally care a lot about conformity, and they have great respect for authority—in word, if not always in deed. They rarely question the legitimacy of authority, even as they learn to maneuver within and around the constraints that adults impose on them.

After puberty, right when Piaget said that children become capable of abstract thought, Kohlberg found that some children begin to think for themselves about the nature of authority, the meaning of justice, and the reasons behind rules and laws. In the two "post-conventional" stages, adolescents still value honesty and respect rules and laws, but now they sometimes justify dishonesty or law-breaking in pursuit of still higher goods, particularly justice. Kohlberg painted an inspiring rationalist image of children as "moral philosophers" trying to work out coherent ethical systems for themselves.[10] In the post-conventional stages, they finally get good at it. Kohlberg's dilemmas were a tool for measuring these dramatic advances in moral reasoning.

THE LIBERAL CONSENSUS

Mark Twain once said that "to a man with a hammer, everything looks like a nail." Once Kohlberg developed his moral dilemmas and his scoring techniques, the psychological community had a new hammer, and a thousand graduate students used it to pound out dissertations on moral reasoning. But there's a deeper reason so many young psychologists began to study morality from a rationalist perspective, and this was Kohlberg's second great innovation: he used his research to build a scientific justification for a secular liberal moral order.

Kohlberg's most influential finding was that the most morally advanced kids (according to his scoring technique) were those who had frequent opportunities for role taking—for putting themselves into another person's shoes and looking at a problem from that per-

son's perspective. Egalitarian relationships (such as with peers) invite role taking, but hierarchical relationships (such as with teachers and parents) do not. It's really hard for a child to see things from the teacher's point of view, because the child has never been a teacher. Piaget and Kohlberg both thought that parents and other authorities were *obstacles* to moral development. If you want your kids to learn about the physical world, let them play with cups and water; don't lecture them about the conservation of volume. And if you want your kids to learn about the social world, let them play with other kids and resolve disputes; don't lecture them about the Ten Commandments. And, for heaven's sake, don't force them to obey God or their teachers or you. That will only freeze them at the conventional level.

Kohlberg's timing was perfect. Just as the first wave of baby boomers was entering graduate school, he transformed moral psychology into a boomer-friendly ode to justice, and he gave them a tool to measure children's progress toward the liberal ideal. For the next twenty-five years, from the 1970s through the 1990s, moral psychologists mostly just interviewed young people about moral dilemmas and analyzed their justifications.[11] Most of this work was not politically motivated—it was careful and honest scientific research. But by using a framework that predefined morality as justice while denigrating authority, hierarchy, and tradition, it was inevitable that the research would support worldviews that were secular, questioning, and egalitarian.

AN EASIER TEST

If you force kids to explain complex notions, such as how to balance competing concerns about rights and justice, you're guaranteed to find age trends because kids get so much more articulate with each passing year. But if you are searching for the first appearance of a moral concept, then you'd better find a technique that doesn't require much verbal skill. Kohlberg's former student Elliot Turiel developed such a technique. His innovation was to tell children short stories about other kids who break rules and then give them a series of sim-

ple yes-or-no probe questions. For example, you tell a story about a child who goes to school wearing regular clothes, even though his school requires students to wear a uniform. You start by getting an overall judgment: "Is that OK, what the boy did?" Most kids say no. You ask if there's a rule about what to wear. ("Yes.") Then you probe to find out what kind of rule it is: "What if the teacher said it was OK for the boy to wear his regular clothes, then would it be OK?" and "What if this happened in another school, where they don't have any rules about uniforms, then would it be OK?"

Turiel discovered that children as young as five usually say that the boy was wrong to break the rule, but that it would be OK if the teacher gave permission or if it happened in another school where there was no such rule. Children recognize that rules about clothing, food, and many other aspects of life are *social conventions*, which are arbitrary and changeable to some extent.[12]

But if you ask kids about actions that hurt other people, such as a girl who pushes a boy off a swing because she wants to use it, you get a very different set of responses. Nearly all kids say that the girl was wrong and that she'd be wrong even if the teacher said it was OK, and even if this happened in another school where there were no rules about pushing kids off swings. Children recognize that rules that prevent harm are *moral rules*, which Turiel defined as rules related to "justice, rights, and welfare pertaining to how people ought to relate to each other."[13]

In other words, young children don't treat all rules the same, as Piaget and Kohlberg had supposed. Kids can't talk like moral philosophers, but they are busy sorting social information in a sophisticated way. They seem to grasp early on that rules that prevent harm are special, important, unalterable, and universal. And this realization, Turiel said, was the foundation of all moral development. Children construct their moral understanding on the bedrock of the absolute moral truth that *harm is wrong*. Specific rules may vary across cultures, but in all of the cultures Turiel examined, children still made a distinction between moral rules and conventional rules.[14]

Turiel's account of moral development differed in many ways from Kohlberg's, but the political implications were similar: morality

is about *treating individuals well*. It's about harm and fairness (not loyalty, respect, duty, piety, patriotism, or tradition). Hierarchy and authority are generally bad things (so it's best to let kids figure things out for themselves). Schools and families should therefore embody progressive principles of equality and autonomy (not authoritarian principles that enable elders to train and constrain children).

MEANWHILE, IN THE REST OF THE WORLD . . .

Kohlberg and Turiel had pretty much defined the field of moral psychology by the time I sat in Jon Baron's office and decided to study morality.[15] The field I entered was vibrant and growing, yet something about it felt wrong to me. It wasn't the politics—I was very liberal back then, twenty-four years old and full of indignation at Ronald Reagan and conservative groups such as the righteously named Moral Majority. No, the problem was that the things I was reading were so . . . dry. I had grown up with two sisters, close in age to me. We fought every day, using every dirty rhetorical trick we could think of. Morality was such a passionate affair in my family, yet the articles I was reading were all about reasoning and cognitive structures and domains of knowledge. It just seemed too cerebral. There was hardly any mention of emotion.

As a first-year graduate student, I didn't have the confidence to trust my instincts, so I forced myself to continue reading. But then, in my second year, I took a course on cultural psychology and was captivated. The course was taught by a brilliant anthropologist, Alan Fiske, who had spent many years in West Africa studying the psychological foundations of social relationships.[16] Fiske asked us all to read several ethnographies (book-length reports of an anthropologist's fieldwork), each of which focused on a different topic, such as kinship, sexuality, or music. But no matter the topic, morality turned out to be a central theme.

I read a book on witchcraft among the Azande of Sudan.[17] It turns out that witchcraft beliefs arise in surprisingly similar forms in many parts of the world, which suggests either that there really are witches

or (more likely) that there's something about human minds that often generates this cultural institution. The Azande believed that witches were just as likely to be men as women, and the fear of being called a witch made the Azande careful not to make their neighbors angry or envious. That was my first hint that groups create supernatural beings not to explain the universe but to order their societies.[18]

I read a book about the Ilongot, a tribe in the Philippines whose young men gained honor by cutting off people's heads.[19] Some of these beheadings were revenge killings, which offered Western readers a motive they could understand. But many of these murders were committed against strangers who were not involved in any kind of feud with the killer. The author explained these most puzzling killings as ways that small groups of men channeled resentments and frictions within the group into a group-strengthening "hunting party," capped off by a long night of communal celebratory singing. This was my first hint that morality often involves tension *within* the group linked to competition *between* different groups.

These ethnographies were fascinating, often beautifully written, and intuitively graspable despite the strangeness of their content. Reading each book was like spending a week in a new country: confusing at first, but gradually you tune up, finding yourself better able to guess what's going to happen next. And as with all foreign travel, you learn as much about where you're from as where you're visiting. I began to see the United States and Western Europe as extraordinary historical exceptions—new societies that had found a way to strip down and thin out the thick, all-encompassing moral orders that the anthropologists wrote about.

Nowhere was this thinning more apparent than in our lack of rules about what the anthropologists call "purity" and "pollution." Contrast us with the Hua of New Guinea, who have developed elaborate networks of food taboos that govern what men and women may eat. In order for their boys to become men, they have to avoid foods that in any way resemble vaginas, including anything that is red, wet, slimy, comes from a hole, or has hair. It sounds at first like arbitrary superstition mixed with the predictable sexism of a patriarchal society. Turiel would call these rules social conventions, because the Hua

don't believe that men in other tribes have to follow these rules. But the Hua certainly seemed to think of their food rules as moral rules. They talked about them constantly, judged each other by their food habits, and governed their lives, duties, and relationships by what the anthropologist Anna Meigs called "a religion of the body."[20]

But it's not just hunter-gatherers in rain forests who believe that bodily practices can be moral practices. When I read the Hebrew Bible, I was shocked to discover how much of the book—one of the sources of Western morality—was taken up with rules about food, menstruation, sex, skin, and the handling of corpses. Some of these rules were clear attempts to avoid disease, such as the long sections of Leviticus on leprosy. But many of the rules seemed to follow a more emotional logic about avoiding disgust. For example, the Bible prohibits Jews from eating or even touching "the swarming things that swarm upon the earth" (and just think how much more disgusting a swarm of mice is than a single mouse).[21] Other rules seemed to follow a conceptual logic involving keeping categories pure or not mixing things together (such as clothing made from two different fibers).[22]

So what's going on here? If Turiel was right that morality is really about harm, then why do most non-Western cultures moralize so many practices that seem to have nothing to do with harm? Why do many Christians and Jews believe that "cleanliness is next to godliness"?[23] And why do so many Westerners, even secular ones, continue to see choices about food and sex as being heavily loaded with moral significance? Liberals sometimes say that religious conservatives are sexual prudes for whom anything other than missionary-position intercourse within marriage is a sin. But conservatives can just as well make fun of liberal struggles to choose a balanced breakfast—balanced among moral concerns about free-range eggs, fair-trade coffee, naturalness, and a variety of toxins, some of which (such as genetically modified corn and soybeans) pose a greater threat spiritually than biologically. Even if Turiel was right that children lock onto harmfulness as a method for identifying immoral actions, I couldn't see how kids in the West—let alone among the Azande, the Ilongot, and the Hua—could have come to all this purity and pollution stuff on their own. There must be more to moral development than kids

constructing rules as they take the perspectives of other people and feel their pain. There must be something beyond rationalism.

THE GREAT DEBATE

When anthropologists wrote about morality, it was as though they spoke a different language from the psychologists I had been reading. The Rosetta stone that helped me translate between the two fields was a paper that had just been published by Fiske's former advisor, Richard Shweder, at the University of Chicago.[24] Shweder is a psychological anthropologist who had lived and worked in Orissa, a state on the east coast of India. He had found large differences in how Oriyans (residents of Orissa) and Americans thought about personality and individuality, and these differences led to corresponding differences in how they thought about morality. Shweder quoted the anthropologist Clifford Geertz on how unusual Westerners are in thinking about people as discrete individuals:

> The Western conception of the person as a bounded, unique, more or less integrated motivational and cognitive universe, a dynamic center of awareness, emotion, judgment, and action organized into a distinctive whole and set contrastively both against other such wholes and against its social and natural background, is, however incorrigible it may seem to us, a rather peculiar idea within the context of the world's cultures.[25]

Shweder offered a simple idea to explain why the self differs so much across cultures: all societies must resolve a small set of questions about how to order society, the most important being how to balance the needs of individuals and groups. There seem to be just two primary ways of answering this question. Most societies have chosen the *sociocentric* answer, placing the needs of groups and institutions first, and subordinating the needs of individuals. In contrast, the *individualistic* answer places individuals at the center and makes society a

servant of the individual.[26] The sociocentric answer dominated most of the ancient world, but the individualistic answer became a powerful rival during the Enlightenment. The individualistic answer largely vanquished the sociocentric approach in the twentieth century as individual rights expanded rapidly, consumer culture spread, and the Western world reacted with horror to the evils perpetrated by the ultrasociocentric fascist and communist empires. (European nations with strong social safety nets are not sociocentric on this definition. They just do a very good job of protecting *individuals* from the vicissitudes of life.)

Shweder thought that the theories of Kohlberg and Turiel were produced by and for people from individualistic cultures. He doubted that those theories would apply in Orissa, where morality was sociocentric, selves were interdependent, and no bright line separated moral rules (preventing harm) from social conventions (regulating behaviors not linked directly to harm). To test his ideas, he and two collaborators came up with thirty-nine very short stories in which someone does something that would violate a rule either in the United States or in Orissa. The researchers then interviewed 180 children (ranging in age from five to thirteen) and 60 adults who lived in Hyde Park (the neighborhood surrounding the University of Chicago) about these stories. They also interviewed a matched sample of Brahmin children and adults in the town of Bhubaneswar (an ancient pilgrimage site in Orissa),[27] and 120 people from low ("untouchable") castes. Altogether it was an enormous undertaking—six hundred long interviews in two very different cities.

The interview used Turiel's method, more or less, but the scenarios covered many more behaviors than Turiel had ever asked about. As you can see in the top third of figure 1.1, people in some of the stories obviously hurt other people or treated them unfairly, and subjects (the people being interviewed) in both countries condemned these actions by saying that they were wrong, unalterably wrong, and universally wrong. But the Indians would not condemn other cases that seemed (to Americans) just as clearly to involve harm and unfairness (see middle third).

Most of the thirty-nine stories portrayed no harm or unfairness,

Actions that Indians and Americans agreed were wrong:

- While walking, a man saw a dog sleeping on the road. He walked up to it and kicked it.
- A father said to his son, "If you do well on the exam, I will buy you a pen." The son did well on the exam, but the father did not give him anything.

Actions that Americans said were wrong but Indians said were acceptable:

- A young married woman went alone to see a movie without informing her husband. When she returned home her husband said, "If you do it again, I will beat you black and blue." She did it again; he beat her black and blue. (Judge the husband.)
- A man had a married son and a married daughter. After his death his son claimed most of the property. His daughter got little. (Judge the son.)

Actions that Indians said were wrong but Americans said were acceptable:

- In a family, a twenty-five-year-old son addresses his father by his first name.
- A woman cooked rice and wanted to eat with her husband and his elder brother. Then she ate with them. (Judge the woman.)
- A widow in your community eats fish two or three times a week.
- After defecation a woman did not change her clothes before cooking.

FIGURE 1.1. *Some of the thirty-nine stories used in Shweder, Mahapatra, and Miller 1987.*

at least none that could have been obvious to a five-year-old child, and nearly all Americans said that these actions were permissible (see the bottom third of figure 1.1). If Indians said that these actions were wrong, then Turiel would predict that they were condemning the actions merely as violations of social conventions. Yet most of the Indian subjects—even the five-year-old children—said that these actions were wrong, universally wrong, and unalterably wrong. Indian practices related to food, sex, clothing, and gender relations were almost always judged to be moral issues, not social conventions, and there were few differences between the adults and children within each city. In other words, Shweder found almost no trace of social conventional thinking in the sociocentric culture of Orissa, where, as he put it, "the social order is a moral order." Morality was much broader and thicker in Orissa; almost any practice could be loaded up with moral force. And if that was true, then Turiel's theory became less plausible. Children were not figuring out morality for themselves, based on the bedrock certainty that harm is bad.

Even in Chicago, Shweder found relatively little evidence of social-conventional thinking. There were plenty of stories that contained no obvious harm or injustice, such as a widow eating fish, and Americans predictably said that those cases were fine. But more important, they didn't see these behaviors as social conventions that could be changed by popular consent. They believed that widows should be able to eat whatever they darn well please, and if there's some other country where people try to limit widows' freedoms, well, they're wrong to do so. Even in the United States the social order is a moral order, but it's an individualistic order built up around the protection of individuals and their freedom. The distinction between morals and mere conventions is not a tool that children everywhere use to self-construct their moral knowledge. Rather, the distinction turns out to be a cultural artifact, a necessary by-product of the individualistic answer to the question of how individuals and groups relate. When you put individuals first, before society, then any rule or social practice that limits personal freedom can be questioned. If it doesn't protect somebody from harm, then it can't be morally justified. It's just a social convention.

Shweder's study was a major attack on the whole rationalist approach, and Turiel didn't take it lying down. He wrote a long rebuttal essay pointing out that many of Shweder's thirty-nine stories were trick questions: they had very different meanings in India and America.[28] For example, Hindus in Orissa believe that fish is a "hot" food that will stimulate a person's sexual appetite. If a widow eats hot foods, she is more likely to have sex with someone, which would offend the spirit of her dead husband and prevent her from reincarnating at a higher level. Turiel argued that once you take into account Indian "informational assumptions" about the way the world works, you see that most of Shweder's thirty-nine stories really *were* moral violations, harming victims in ways that Americans could not see. So Shweder's study didn't contradict Turiel's claims; it might even support them, if we could find out for sure whether Shweder's Indian subjects saw harm in the stories.

DISGUST AND DISRESPECT

When I read the Shweder and Turiel essays, I had two strong reactions. The first was an intellectual agreement with Turiel's defense. Shweder had used "trick" questions not to be devious but to demonstrate that rules about food, clothing, ways of addressing people, and other seemingly conventional matters could all get woven into a thick moral web. Nonetheless, I agreed with Turiel that Shweder's study was missing an important experimental control: he didn't ask his subjects about harm. If Shweder wanted to show that morality extended beyond harm in Orissa, he had to show that people were willing to morally condemn actions that *they themselves* stated were harmless.

My second reaction was a gut feeling that Shweder was ultimately right. His explanation of sociocentric morality fit so perfectly with the ethnographies I had read in Fiske's class. His emphasis on the moral emotions was so satisfying after reading all that cerebral cognitive-developmental work. I thought that if somebody ran the right study—one that controlled for perceptions of harm—Shweder's

claims about cultural differences would survive the test. I spent the next semester figuring out how to become that somebody.

I started writing very short stories about people who do offensive things, but do them in such a way that nobody is harmed. I called these stories "harmless taboo violations," and you read two of them at the start of this chapter (about dog-eating and chicken- . . . eating). I made up dozens of these stories but quickly found that the ones that worked best fell into two categories: disgust and disrespect. If you want to give people a quick flash of revulsion but deprive them of any victim they can use to justify moral condemnation, ask them about people who do disgusting or disrespectful things, but make sure the actions are done in private so that nobody else is offended. For example, one of my disrespect stories was: "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."

My idea was to give adults and children stories that pitted gut feelings about important cultural norms against reasoning about harmlessness, and then see which force was stronger. Turiel's rationalism predicted that reasoning about harm is the basis of moral judgment, so even though people might say it's wrong to eat your dog, they would have to treat the act as a violation of a social convention. (*We* don't eat our dogs, but hey, if people in another country want to eat their ex-pets rather than bury them, who are we to criticize?) Shweder's theory, on the other hand, said that Turiel's predictions should hold among members of individualistic secular societies but not elsewhere. I now had a study designed. I just had to find the elsewhere.

I spoke Spanish fairly well, so when I learned that a major conference of Latin American psychologists was to be held in Buenos Aires in July 1989, I bought a plane ticket. I had no contacts and no idea how to start an international research collaboration, so I just went to every talk that had anything to do with morality. I was chagrined to discover that psychology in Latin America was not very scientific. It was heavily theoretical, and much of that theory was Marxist, focused on oppression, colonialism, and power. I was beginning to despair

when I chanced upon a session run by some Brazilian psychologists who were using Kohlbergian methods to study moral development. I spoke afterward to the chair of the session, Angela Biaggio, and her graduate student Silvia Koller. Even though they both liked Kohlberg's approach, they were interested in hearing about alternatives. Biaggio invited me to visit them after the conference at their university in Porto Alegre, the capital of the southernmost state in Brazil.

Southern Brazil is the most European part of the country, settled largely by Portuguese, German, and Italian immigrants in the nineteenth century. With its modern architecture and middle-class prosperity, Porto Alegre didn't look anything like the Latin America of my imagination, so at first I was disappointed. I wanted my cross-cultural study to involve someplace exotic, like Orissa. But Silvia Koller was a wonderful collaborator, and she had two great ideas about how to increase our cultural diversity. First, she suggested we run the study across social class. The divide between rich and poor is so vast in Brazil that it's as though people live in different countries. We decided to interview adults and children from the educated middle class, and also from the lower class—adults who worked as servants for wealthy people (and who rarely had more than an eighth-grade education) and children from a public school in the neighborhood where many of the servants lived. Second, Silvia had a friend who had just been hired as a professor in Recife, a city in the northeastern tip of the country, a region that is culturally very different from Porto Alegre. Silvia arranged for me to visit her friend, Graça Dias, the following month.

Silvia and I worked for two weeks with a team of undergraduate students, translating the harmless taboo stories into Portuguese, selecting the best ones, refining the probe questions, and testing our interview script to make sure that everything was understandable, even by the least educated subjects, some of whom were illiterate. Then I went off to Recife, where Graça and I trained a team of students to conduct interviews in exactly the way they were being done in Porto Alegre. In Recife I finally felt like I was working in an exotic tropical locale, with Brazilian music wafting through the streets and ripe mangoes falling from the trees. More important, the people of

I once overheard a Kohlberg-style moral judgment interview being conducted in the bathroom of a McDonald's restaurant in northern Indiana. The person interviewed—the subject—was a Caucasian male roughly thirty years old. The interviewer was a Caucasian male approximately four years old. The interview began at adjacent urinals:

> INTERVIEWER: *Dad, what would happen if I pooped in here [the urinal]?*
> SUBJECT: *It would be yucky. Go ahead and flush. Come on, let's go wash our hands.*
> *[The pair then moved over to the sinks]*
> INTERVIEWER: *Dad, what would happen if I pooped in the sink?*
> SUBJECT: *The people who work here would get mad at you.*
> INTERVIEWER: *What would happen if I pooped in the sink at home?*
> SUBJECT: *I'd get mad at you.*
> INTERVIEWER: *What would happen if you pooped in the sink at home?*
> SUBJECT: *Mom would get mad at me.*
> INTERVIEWER: *Well, what would happen if we all pooped in the sink at home?*
> SUBJECT: *[pause] I guess we'd all get in trouble.*
> INTERVIEWER: *[laughing] Yeah, we'd all get in trouble!*
> SUBJECT: *Come on, let's dry our hands. We have to go.*

Note the skill and persistence of the interviewer, who probes for a deeper answer by changing the transgression to remove the punisher. Yet even when everyone cooperates in the rule violation so that nobody can play the role of punisher, the subject still clings to a notion of cosmic justice in which, somehow, the whole family would "get in trouble."

Of course, the father is not really trying to demonstrate his best moral reasoning. Moral reasoning is usually done to influence other people (see chapter 4), and what the father is trying to do is get his curious son to feel the right emotions—disgust and fear—to motivate appropriate bathroom behavior.

INVENTING VICTIMS

Even though the results came out just as Shweder had predicted, there were a number of surprises along the way. The biggest surprise was that so many subjects tried to invent victims. I had written the stories carefully to remove all conceivable harm to other people, yet in 38 percent of the 1,620 times that people heard a harmless-offensive story, they claimed that somebody was harmed. In the dog story, for example, many people said that the family itself would be harmed because they would get sick from eating dog meat. Was this an example of the "informational assumptions" that Turiel had talked about? Were people really condemning the actions *because* they foresaw these harms, or was it the reverse process—were people *inventing* these harms because they had already condemned the actions?

I conducted many of the Philadelphia interviews myself, and it was obvious that most of these supposed harms were post hoc fabrications. People usually condemned the actions very quickly—they didn't seem to need much time to decide what they thought. But it often took them a while to come up with a victim, and they usually offered those victims up halfheartedly and almost apologetically. As one subject said, "Well, I don't know, maybe the woman will feel guilty afterward about throwing out her flag?" Many of these victim claims were downright preposterous, such as the child who justified his condemnation of the flag shredder by saying that the rags might clog up the toilet and cause it to overflow.

But something even more interesting happened when I or the other interviewers challenged these invented-victim claims. I had trained my interviewers to correct people gently when they made claims that contradicted the text of the story. For example, if someone said, "It's wrong to cut up the flag because a neighbor might see her do it, and he might be offended," the interviewer replied, "Well, it says here in the story that nobody saw her do it. So would you still say it was wrong for her to cut up her flag?" Yet even when subjects recognized that their victim claims were bogus, they still refused to say that the act was OK. Instead, they kept searching for another victim. They

said things like "I know it's wrong, but I just can't think of a reason why." They seemed to be *morally dumbfounded*—rendered speechless by their inability to explain verbally what they knew intuitively.[29]

These subjects were reasoning. They were working quite hard at reasoning. But it was not reasoning in search of truth; it was reasoning in support of their emotional reactions. It was reasoning as described by the philosopher David Hume, who wrote in 1739 that "reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them."[30]

I had found evidence for Hume's claim. I had found that moral reasoning was often a servant of moral emotions, and this was a challenge to the rationalist approach that dominated moral psychology. I published these findings in one of the top psychology journals in October 1993[31] and then waited nervously for the response. I knew that the field of moral psychology was not going to change overnight just because one grad student produced some data that didn't fit into the prevailing paradigm. I knew that debates in moral psychology could be quite heated (though always civil). What I did not expect, however, was that there would be no response at all. Here I thought I had done the definitive study to settle a major debate in moral psychology, yet almost nobody cited my work—not even to attack it—in the first five years after I published it.

My dissertation landed with a silent thud in part because I published it in a social psychology journal. But in the early 1990s, the field of moral psychology was still a part of developmental psychology. If you called yourself a moral psychologist, it meant that you studied moral reasoning and how it changed with age, and you cited Kohlberg extensively whether you agreed with him or not.

But psychology itself was about to change and become a lot more emotional.

IN SUM

Where does morality come from? The two most common answers have long been that it is innate (the nativist answer) or that it comes

from childhood learning (the empiricist answer). In this chapter I considered a third possibility, the rationalist answer, which dominated moral psychology when I entered the field: that morality is self-constructed by children on the basis of their experiences with harm. Kids know that harm is wrong because they hate to be harmed, and they gradually come to see that it is therefore wrong to harm others, which leads them to understand fairness and eventually justice. I explained why I came to reject this answer after conducting research in Brazil and the United States. I concluded instead that:

- The moral domain varies by culture. It is unusually narrow in Western, educated, and individualistic cultures. Sociocentric cultures broaden the moral domain to encompass and regulate more aspects of life.
- People sometimes have gut feelings—particularly about disgust and disrespect—that can drive their reasoning. Moral reasoning is sometimes a post hoc fabrication.
- Morality can't be entirely self-constructed by children based on their growing understanding of harm. Cultural learning or guidance must play a larger role than rationalist theories had given it.

If morality doesn't come primarily from reasoning, then that leaves some combination of innateness and social learning as the most likely candidates. In the rest of this book I'll try to explain how morality can be innate (as a set of evolved intuitions) and learned (as children learn to apply those intuitions within a particular culture). We're born to be righteous, but we have to learn what, exactly, people like us should be righteous about.

The Intuitive Dog and Its Rational Tail

One of the greatest truths in psychology is that the mind is divided into parts that sometimes conflict.[1] To be human is to feel pulled in different directions, and to marvel—sometimes in horror—at your inability to control your own actions. The Roman poet Ovid lived at a time when people thought diseases were caused by imbalances of bile, but he knew enough psychology to have one of his characters lament: "I am dragged along by a strange new force. Desire and reason are pulling in different directions. I see the right way and approve it, but follow the wrong."[2]

Ancient thinkers gave us many metaphors to understand this conflict, but few are more colorful than the one in Plato's dialogue *Timaeus*. The narrator, Timaeus, explains how the gods created the universe, including us. Timaeus says that a creator god who was perfect and created only perfect things was filling his new universe with souls—and what could be more perfect in a soul than perfect rationality? So after making a large number of perfect, rational souls, the creator god decided to take a break, delegating the last bits of creation to some lesser deities, who did their best to design vessels for these souls.

The deities began by encasing the souls in that most perfect of shapes, the sphere, which explains why our heads are more or less round. But they quickly realized that these spherical heads would face

When nature assigned us the same habitation, she gave us over it a divided empire. To you she allotted the field of science; to me that of morals. When the circle is to be squared, or the orbit of a comet to be traced; when the arch of greatest strength, or the solid of least resistance is to be investigated, take up the problem; it is yours; nature has given me no cognizance of it. In like manner, in denying to you the feelings of sympathy, of benevolence, of gratitude, of justice, of love, of friendship, she has excluded you from their control. To these she has adapted the mechanism of the heart. Morals were too essential to the happiness of man to be risked on the incertain combinations of the head. She laid their foundation therefore in sentiment, not in science.[8]

So now we have three models of the mind. Plato said that reason *ought* to be the master, even if philosophers are the only ones who can reach a high level of mastery.[9] Hume said that reason is and ought to be the servant of the passions. And Jefferson gives us a third option, in which reason and sentiment are (and ought to be) independent co-rulers, like the emperors of Rome, who divided the empire into eastern and western halves. Who is right?

WILSON'S PROPHECY

Plato, Hume, and Jefferson tried to understand the design of the human mind without the help of the most powerful tool ever devised for understanding the design of living things: Darwin's theory of evolution. Darwin was fascinated by morality because any example of cooperation among living creatures had to be squared with his general emphasis on competition and the "survival of the fittest."[10] Darwin offered several explanations for how morality could have evolved, and many of them pointed to emotions such as sympathy, which he thought was the "foundation-stone" of the social instincts.[11] He also wrote about feelings of shame and pride, which were associated with

the desire for a good reputation. Darwin was a nativist about morality: he thought that natural selection gave us minds that were preloaded with moral emotions.

But as the social sciences advanced in the twentieth century, their course was altered by two waves of moralism that turned nativism into a moral offense. The first was the horror among anthropologists and others at "social Darwinism"—the idea (raised but not endorsed by Darwin) that the richest and most successful nations, races, and individuals are the fittest. Therefore, giving charity to the poor interferes with the natural progress of evolution: it allows the poor to breed.[12] The claim that some races were innately superior to others was later championed by Hitler, and so if Hitler was a nativist, then all nativists were Nazis. (That conclusion is illogical, but it makes sense emotionally if you dislike nativism.)[13]

The second wave of moralism was the radical politics that washed over universities in America, Europe, and Latin America in the 1960s and 1970s. Radical reformers usually want to believe that human nature is a blank slate on which any utopian vision can be sketched. If evolution gave men and women different sets of desires and skills, for example, that would be an obstacle to achieving gender equality in many professions. If nativism could be used to justify existing power structures, then nativism must be wrong. (Again, this is a logical error, but this is the way righteous minds work.)

The cognitive scientist Steven Pinker was a graduate student at Harvard in the 1970s. In his 2002 book *The Blank Slate: The Modern Denial of Human Nature,* Pinker describes the ways scientists betrayed the values of science to maintain loyalty to the progressive movement. Scientists became "moral exhibitionists" in the lecture hall as they demonized fellow scientists and urged their students to evaluate ideas not for their truth but for their consistency with progressive ideals such as racial and gender equality.[14]

Nowhere was the betrayal of science more evident than in the attacks on Edward O. Wilson, a lifelong student of ants and ecosystems. In 1975 Wilson published *Sociobiology: The New Synthesis.* The book explored how natural selection, which indisputably shaped animal bodies, also shaped animal behavior. That wasn't controver-

sial, but Wilson had the audacity to suggest in his final chapter that natural selection also influenced *human* behavior. Wilson believed that there is such a thing as human nature, and that human nature constrains the range of what we can achieve when raising our children or designing new social institutions.

Wilson used ethics to illustrate his point. He was a professor at Harvard, along with Lawrence Kohlberg and the philosopher John Rawls, so he was well acquainted with their brand of rationalist theorizing about rights and justice.[15] It seemed clear to Wilson that what the rationalists were *really* doing was generating clever justifications for moral intuitions that were best explained by evolution. Do people believe in human rights because such rights actually exist, like mathematical truths, sitting on a cosmic shelf next to the Pythagorean theorem just waiting to be discovered by Platonic reasoners? Or do people feel revulsion and sympathy when they read accounts of torture, and then invent a story about universal rights to help justify their feelings?

Wilson sided with Hume. He charged that what moral philosophers were really doing was fabricating justifications after "consulting the emotive centers" of their own brains.[16] He predicted that the study of ethics would soon be taken out of the hands of philosophers and "biologicized," or made to fit with the emerging science of human nature. Such a linkage of philosophy, biology, and evolution would be an example of the "new synthesis" that Wilson dreamed of, and that he later referred to as *consilience*—the "jumping together" of ideas to create a unified body of knowledge.[17]

Prophets challenge the status quo, often earning the hatred of those in power. Wilson therefore deserves to be called a prophet of moral psychology. He was harassed and excoriated in print and in public.[18] He was called a fascist, which justified (for some) the charge that he was a racist, which justified (for some) the attempt to stop him from speaking in public. Protesters who tried to disrupt one of his scientific talks rushed the stage and chanted, "Racist Wilson, you can't hide, we charge you with genocide."[19]

THE EMOTIONAL NINETIES

By the time I entered graduate school, in 1987, the shooting had stopped and sociobiology had been discredited—at least, that's the message I picked up from hearing scientists use the word as a pejorative term for the naive attempt to reduce psychology to evolution. Moral psychology was not about evolved emotions, it was about the development of reasoning and information processing.[20]

Yet when I looked outside of psychology, I found many wonderful books on the emotional basis of morality. I read Frans de Waal's *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*.[21] De Waal did not claim that chimpanzees had morality; he argued only that chimps (and other apes) have most of the psychological building blocks that humans use to construct moral systems and communities. These building blocks are largely emotional, such as feelings of sympathy, fear, anger, and affection.

I also read *Descartes' Error*, by the neuroscientist Antonio Damasio.[22] Damasio had noticed an unusual pattern of symptoms in patients who had suffered brain damage to a specific part of the brain—the ventromedial (i.e., bottom-middle) prefrontal cortex (abbreviated vmPFC; it's the region just behind and above the bridge of the nose). Their emotionality dropped nearly to zero. They could look at the most joyous or gruesome photographs and feel nothing. They retained full knowledge of what was right and wrong, and they showed no deficits in IQ. They even scored well on Kohlberg's tests of moral reasoning. Yet when it came to making decisions in their personal lives and at work, they made foolish decisions or no decisions at all. They alienated their families and their employers, and their lives fell apart.

Damasio's interpretation was that gut feelings and bodily reactions were *necessary* to think rationally, and that one job of the vmPFC was to integrate those gut feelings into a person's conscious deliberations. When you weigh the advantages and disadvantages of murdering your parents . . . you can't even do it, because feelings of horror come rushing in through the vmPFC.

But Damasio's patients could think about anything, with no filtering or coloring from their emotions. With the vmPFC shut down, every option at every moment felt as good as every other. The only way to make a decision was to examine each option, weighing the pros and cons using conscious, verbal reasoning. If you've ever shopped for an appliance about which you have few feelings—say, a washing machine—you know how hard it can be once the number of options exceeds six or seven (which is the capacity of our short-term memory). Just imagine what your life would be like if at every moment, in every social situation, picking the right thing to do or say became like picking the best washing machine among ten options, minute after minute, day after day. You'd make foolish decisions too.

Damasio's findings were as anti-Platonic as could be. Here were people in whom brain damage had essentially shut down communication between the rational soul and the seething passions of the body (which, unbeknownst to Plato, were not based in the heart and stomach but in the emotion areas of the brain). No more of those "dreadful but necessary disturbances," those "foolish counselors" leading the rational soul astray. Yet the result of the separation was not the liberation of reason from the thrall of the passions. It was the shocking revelation that reasoning *requires* the passions. Jefferson's model fits better: when one co-emperor is knocked out and the other tries to rule the empire by himself, he's not up to the task.

If Jefferson's model were correct, however, then Damasio's patients should still have fared well in the half of life that was always ruled by the head. Yet the collapse of decision making, even in purely analytic and organizational tasks, was pervasive. The head can't even do head stuff without the heart. So Hume's model fit these cases best: when the master (passions) drops dead, the servant (reasoning) has neither the ability nor the desire to keep the estate running. Everything goes to ruin.

enough, Scott couldn't whip up any dumbfounding with the Heinz story. People offered good reasons for their answers, and Scott was not able to get them to abandon principles such as "Life is more important than property."

We also chose two scenarios that played more directly on gut feelings. In the "roach juice" scenario, Scott opened a small can of apple juice, poured it into a new plastic cup, and asked the subject to take a sip. Everyone did. Then Scott brought out a white plastic box and said:

> I have here in this container a sterilized cockroach. We bought some cockroaches from a laboratory supply company. The roaches were raised in a clean environment. But just to be certain, we sterilized the roaches again in an autoclave, which heats everything so hot that no germs can survive. I'm going to dip this cockroach into the juice, like this [using a tea strainer]. Now, would you take a sip?

In the second scenario, Scott offered subjects $2 if they would sign a piece of paper that said: *I, _____, hereby sell my soul, after my death, to Scott Murphy, for the sum of $2.* There was a line for a signature, and below the line was this note: *This form is part of a psychology experiment. It is* NOT *a legal or binding contract, in any way.*[27] Scott also told them they could rip up the paper as soon as they signed it, and they'd still get their $2.

Only 23 percent of subjects were willing to sign the paper without any goading from Scott. We were a bit surprised to find that 37 percent were willing to take a sip of the roach juice.[28] In these cases, Scott couldn't play devil's advocate.

For the majorities who said no, however, Scott asked them to explain their reasons and did his best to challenge those reasons. Scott convinced an extra 10 percent to sip the juice, and an extra 17 percent to sign the soul-selling paper. But most people in both scenarios clung to their initial refusal, even though many of them could not generate good reasons. A few people confessed that they

were atheists, didn't believe in souls, and yet still felt uncomfortable about signing.

Here too there wasn't much dumbfounding. People felt that it was ultimately their own choice whether or not to drink the juice or sign the paper, so most subjects seemed comfortable saying, "I just don't want to do it, even though I can't give you a reason."

The main point of the study was to examine responses to two harmless taboo violations. We wanted to know if the moral judgment of disturbing but harmless events would look more like judgments in the Heinz task (closely linked to reasoning), or like those in the roach juice and soul-selling tasks (where people readily confessed that they were following gut feelings). Here's one story we used:

> Julie and Mark, who are sister and brother, are traveling together in France. They are both on summer vacation from college. One night they are staying alone in a cabin near the beach. They decide that it would be interesting and fun if they tried making love. At the very least it would be a new experience for each of them. Julie is already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy it, but they decide not to do it again. They keep that night as a special secret between them, which makes them feel even closer to each other. So what do you think about this? Was it wrong for them to have sex?

In the other harmless-taboo story, Jennifer works in a hospital pathology lab. She's a vegetarian for moral reasons—she think it's wrong to kill animals. But one night she has to incinerate a fresh human cadaver, and she thinks it's a waste to throw away perfectly edible flesh. So she cuts off a piece of flesh and takes it home. Then she cooks it and eats it.

We knew these stories were disgusting, and we expected that they'd trigger immediate moral condemnation. Only 20 percent of subjects said it was OK for Julie and Mark to have sex, and only 13 percent said it was OK for Jennifer to eat part of a cadaver. But

when Scott asked people to explain their judgments and then challenged those explanations, he found exactly the Humean pattern that we had predicted. In these harmless-taboo scenarios, people generated far more reasons and discarded far more reasons than in any of the other scenarios. They seemed to be flailing around, throwing out reason after reason, and rarely changing their minds when Scott proved that their latest reason was not relevant. Here is the transcript of one interview about the incest story:

> EXPERIMENTER: So what do you think about this, was it wrong for Julie and Mark to have sex?
>
> SUBJECT: Yeah, I think it's totally wrong to have sex. You know, because I'm pretty religious and I just think incest is wrong anyway. But, I don't know.
>
> EXPERIMENTER: What's wrong with incest, would you say?
>
> SUBJECT: Um, the whole idea of, well, I've heard— I don't even know if this is true, but in the case, if the girl did get pregnant, the kids become deformed, most of the time, in cases like that.
>
> EXPERIMENTER: But they used a condom and birth control pills—
>
> SUBJECT: Oh, OK. Yeah, you did say that.
>
> EXPERIMENTER: —so there's no way they're going to have a kid.
>
> SUBJECT: Well, I guess the safest sex is abstinence, but, um, uh . . . um, I don't know, I just think that's wrong. I don't know, what did you ask me?
>
> EXPERIMENTER: Was it wrong for them to have sex?
>
> SUBJECT: Yeah, I think it's wrong.
>
> EXPERIMENTER: And I'm trying to find out why, what you think is wrong with it.
>
> SUBJECT: OK, um . . . well . . . let's see, let me think about this. Um—how old were they?
>
> EXPERIMENTER: They were college age, around 20 or so.
>
> SUBJECT: Oh, oh [looks disappointed]. I don't know,

I just . . . it's just not something you're brought up to do. It's just not—well, I mean I wasn't. I assume most people aren't [laughs]. I just think that you shouldn't—I don't—I guess my reason is, um . . . just that, um . . . you're not brought up to it. You don't see it. It's not, um—I don't think it's accepted. That's pretty much it.

EXPERIMENTER: You wouldn't say anything you're not brought up to see is wrong, would you? For example, if you're not brought up to see women working outside the home, would you say that makes it wrong for women to work?

SUBJECT: Um . . . well . . . oh, gosh. This is hard. I really—um, I mean, there's just no way I could change my mind but I just don't know how to—how to show what I'm feeling, what I feel about it. It's crazy![29]

In this transcript and in many others, it's obvious that people were making a moral judgment immediately and emotionally. Reasoning was merely the servant of the passions, and when the servant failed to find any good arguments, the master did not change his mind. We quantified some of the behaviors that seemed most indicative of being morally dumbfounded, and these analyses showed big differences between the way people responded to the harmless-taboo scenarios compared to the Heinz dilemma.[30]

These results supported Hume, not Jefferson or Plato. People made moral judgments quickly and emotionally. Moral reasoning was mostly just a post hoc search for reasons to justify the judgments people had already made. But were these judgments representative of moral judgment in general? I had to write some bizarre stories to give people these flashes of moral intuition that they could not easily explain. That can't be how most of our thinking works, can it?

"SEEING-THAT" VERSUS "REASONING-WHY"

Two years before Scott and I ran the dumbfounding studies I read an extraordinary book that psychologists rarely mention: *Patterns, Thinking, and Cognition*, by Howard Margolis, a professor of public policy at the University of Chicago. Margolis was trying to understand why people's beliefs about political issues are often so poorly connected to objective facts, and he hoped that cognitive science could solve the puzzle. Yet Margolis was turned off by the approaches to thinking that were prevalent in the 1980s, most of which used the metaphor of the mind as a computer.

Margolis thought that a better model for studying higher cognition, such as political thinking, was lower cognition, such as vision, which works largely by rapid unconscious pattern matching. He began his book with an investigation of perceptual illusions, such as the well-known Muller-Lyer illusion (figure 2.2), in which one line continues to look longer than the other even after you know that the two lines are the same length. He then moved on to logic problems such as the Wason 4-card task, in which you are shown four cards on a table.[31] You know that each card comes from a deck in which all cards have a letter on one side and a number on the other. Your task is to choose the smallest number of cards in figure 2.3 that you must turn over to decide whether this rule is true: "If there is a vowel on one side, then there is an even number on the other side."

Everyone immediately sees that you have to turn over the E, but many people also say you need to turn over the 4. They seem
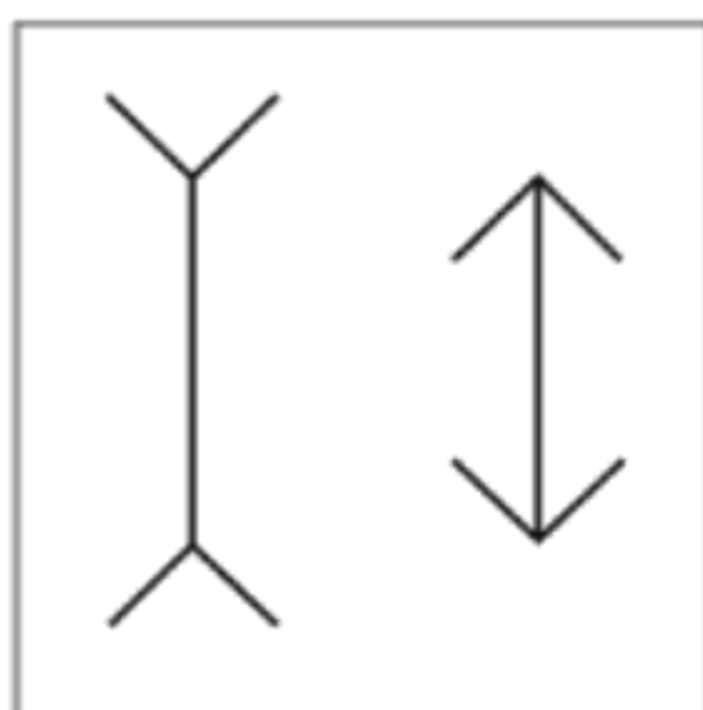

FIGURE 2.2. *The Muller-Lyer illusion.*

steal the drug (his wife's life is at stake), but in this case it's easy to find reasons. Kohlberg had constructed the dilemma to make good reasons available on both sides, so nobody gets dumbfounded.

The roach juice and soul-selling dilemmas instantly make people "see that" they want to refuse, but they don't feel much conversational pressure to offer reasons. Not wanting to drink roach-tainted juice isn't a moral judgment, it's a personal preference. Saying "Because I don't want to" is a perfectly acceptable justification for one's subjective preferences. Yet moral judgments are *not* subjective statements; they are claims that somebody did something wrong. I can't call for the community to punish you simply because I don't like what you're doing. I have to point to something outside of my own preferences, and that pointing is our moral reasoning. We do moral reasoning not to reconstruct the actual reasons why *we ourselves* came to a judgment; we reason to find the best possible reasons why *somebody else ought to join us* in our judgment.[36]

THE RIDER AND THE ELEPHANT

It took me years to appreciate fully the implications of Margolis's ideas. Part of the problem was that my thinking was entrenched in a prevalent but useless dichotomy between cognition and emotion. After failing repeatedly to get cognition to act independently of emotion, I began to realize that the dichotomy made no sense. Cognition just refers to information processing, which includes higher cognition (such as conscious reasoning) as well as lower cognition (such as visual perception and memory retrieval).[37]

Emotion is a bit harder to define. Emotions were long thought to be dumb and visceral, but beginning in the 1980s, scientists increasingly recognized that emotions were filled with cognition. Emotions occur in steps, the first of which is to appraise something that just happened based on whether it advanced or hindered your goals.[38] These appraisals are a kind of information processing; they are cognitions. When an appraisal program detects particular input patterns, it launches a set of changes in your brain that prepare you to

respond appropriately. For example, if you hear someone running up behind you on a dark street, your fear system detects a threat and triggers your sympathetic nervous system, firing up the fight-or-flight response, cranking up your heart rate, and widening your pupils to help you take in more information.

Emotions are not dumb. Damasio's patients made terrible decisions because they were deprived of emotional input into their decision making. *Emotions are a kind of information processing.*[39] Contrasting emotion with cognition is therefore as pointless as contrasting rain with weather, or cars with vehicles.

Margolis helped me ditch the emotion-cognition contrast. His work helped me see that *moral judgment is a cognitive process,* as are all forms of judgment. The crucial distinction is really between *two different kinds of cognition:* intuition and reasoning. Moral emotions are one type of moral intuition, but most moral intuitions are more subtle; they don't rise to the level of emotions.[40] The next time you read a newspaper or drive a car, notice the many tiny flashes of condemnation that flit through your consciousness. Is each such flash an emotion? Or ask yourself whether it is better to save the lives of five strangers or one (assuming all else is equal). Do you need an emotion to tell you to go for the five? Do you need reasoning? No, you just see, instantly, that five is better than one. *Intuition* is the best word to describe the dozens or hundreds of rapid, effortless moral judgments and decisions that we all make every day. Only a few of these intuitions come to us embedded in full-blown emotions.

In *The Happiness Hypothesis,* I called these two kinds of cognition the rider (controlled processes, including "reasoning-why") and the elephant (automatic processes, including emotion, intuition, and all forms of "seeing-that").[41] I chose an elephant rather than a horse because elephants are so much bigger—and smarter—than horses. Automatic processes run the human mind, just as they have been running animal minds for 500 million years, so they're very good at what they do, like software that has been improved through thousands of product cycles. When human beings evolved the capacity for language and reasoning at some point in the last million years, the brain did not rewire itself to hand over the reins to a new and