

ch. 4

VOTE  
FOR  
ME

*The Righteous Mind*

Why Good People Are Divided by  
Politics and Religion

JONATHAN HAIDT



PANTHEON BOOKS, NEW YORK

Copyright © 2012 by Jonathan Haidt

All rights reserved. Published in the United States by Pantheon Books,

a division of Random House, Inc., New York, and in Canada

by Random House of Canada Limited, Toronto.

Pantheon Books and colophon are registered trademarks of Random House, Inc.

Library of Congress Cataloging-in-Publication Data

Haidt, Jonathan.

The righteous mind : why good people are divided

by politics and religion / Jonathan Haidt.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-307-37790-6

1. Ethics. 2. Social psychology. 3. Political psychology. 4. Psychology, Religious.

I. Title.

BJ45.H25 2012 201.615—dc23 2011032036

[www.pantheonbooks.com](http://www.pantheonbooks.com)

[www.righteousmind.com](http://www.righteousmind.com)

Jacket design by Sagmeister Inc.

Printed in the United States of America

First Edition

12 14 16 18 20 19 17 15 13 11

*In memory of my father,  
Harold Haidt*

the other things that would make him like a god among humans. Rather his actions would be in no way different from those of an unjust person, and both would follow the same path.<sup>1</sup>

Glaucon's thought experiment implies that people are only virtuous because they fear the consequences of getting caught—especially the damage to their reputations. Glaucon says he will not be satisfied until Socrates can prove that a just man with a bad reputation is happier than an unjust man who is widely thought to be good.<sup>2</sup>

It's quite a challenge, and Socrates approaches it with an analogy: Justice in a man is like justice in a city (a *polis*, or city-state). He then argues that a just city is one in which there is harmony, cooperation, and a division of labor between all the castes.<sup>3</sup> Farmers farm, carpenters build, and rulers rule. All contribute to the common good, and all lament when misfortune happens to any of them.

But in an unjust city, one group's gain is another's loss, faction schemes against faction, the powerful exploit the weak, and the city is divided against itself. To make sure the *polis* doesn't descend into the chaos of ruthless self-interest, Socrates says that philosophers must rule, for only they will pursue what is truly good, not just what is good for themselves.<sup>4</sup>

Having gotten his listeners to agree to this picture of a just, harmonious, and happy city, Socrates then argues that exactly these sorts of relationships apply within a just, harmonious, and happy *person*. If philosophers must rule the happy city, then reason must rule the happy person. And if reason rules, then it cares about what is truly good, not just about the appearance of virtue.

Plato (who had been a student of Socrates) had a coherent set of beliefs about human nature, and at the core of these beliefs was his faith in the perfectibility of reason. Reason is our original nature, he thought; it was given to us by the gods and installed in our spherical heads. Passions often corrupt reason, but if we can learn to control those passions, our God-given rationality will shine forth and guide us to do the right thing, not the popular thing.

As is often the case in moral philosophy, arguments about what

## FOUR

## Vote for Me (Here's Why)

Suppose the gods were to flip a coin on the day of your birth. Heads, you will be a supremely honest and fair person throughout your life, yet everyone around you will believe you're a scoundrel. Tails, you will cheat and lie whenever it suits your needs, yet everyone around you will believe you're a paragon of virtue. Which outcome would you prefer? Plato's *Republic*—one of the most influential works in the Western canon—is an extended argument that you should pick heads, for your own good. It is better to *be* than to *seem* virtuous.

Early in *The Republic*, Glaucon (Plato's brother) challenges Socrates to prove that justice itself—and not merely the reputation for justice—leads to happiness. Glaucon asks Socrates to imagine what would happen to a man who had the mythical ring of Gyges, a gold ring that makes its wearer invisible at will:

Now, no one, it seems, would be so incorruptible that he would stay on the path of justice or stay away from other people's property, when he could take whatever he wanted from the marketplace with impunity, go into people's houses and have sex with anyone he wished, kill or release from prison anyone he wished, and do all

we *ought* to do depend upon assumptions—often unstated—about human nature and human psychology.<sup>5</sup> And for Plato, the assumed psychology is just plain wrong. In this chapter I'll show that reason is not fit to rule; it was designed to seek justification, not truth. I'll show that Glaucon was right: people care a great deal more about appearance and reputation than about reality. In fact, I'll praise Glaucon for the rest of the book as the guy who got it right—the guy who realized that the most important principle for designing an ethical society is to *make sure that everyone's reputation is on the line all the time*, so that bad behavior will always bring bad consequences.

William James, one of the founders of American psychology, urged psychologists to take a “functionalist” approach to the mind. That means examining things in terms of what they *do*, within a larger system. The function of the heart is to pump blood within the circulatory system, and you can't understand the heart unless you keep that in mind. James applied the same logic to psychology: if you want to understand any mental mechanism or process, you have to know its function within some larger system. Thinking is for doing, he said.<sup>6</sup>

What, then, is the function of moral reasoning? Does it seem to have been shaped, tuned, and crafted (by natural selection) to help us find the truth, so that we can know the right way to behave and condemn those who behave wrongly? If you believe that, then you are a rationalist, like Plato, Socrates, and Kohlberg.<sup>7</sup> Or does moral reasoning seem to have been shaped, tuned, and crafted to help us pursue socially strategic goals, such as guarding our reputations and convincing other people to support us, or our team, in disputes? If you believe that, then you are a Glauconian.

#### WE ARE ALL INTUITIVE POLITICIANS

If you see one hundred insects working together toward a common goal, it's a sure bet they're siblings. But when you see one hundred people working on a construction site or marching off to war, you'd be astonished if they all turned out to be members of one large fam-

ily. Human beings are the world champions of cooperation beyond kinship, and we do it in large part by creating systems of formal and informal accountability. We're really good at holding others accountable for their actions, and we're really skilled at navigating through a world in which others hold us accountable for our own.

Phil Tetlock, a leading researcher in the study of accountability, defines accountability as the “explicit expectation that one will be called upon to justify one's beliefs, feelings, or actions to others,” coupled with an expectation that people will reward or punish us based on how well we justify ourselves.<sup>8</sup> When nobody is answerable to anybody, when slackers and cheaters go unpunished, everything falls apart. (How zealously people punish slackers and cheaters will emerge in later chapters as an important difference between liberals and conservatives.)

Tetlock suggests a useful metaphor for understanding how people behave within the webs of accountability that constitute human societies: we act like *intuitive politicians* striving to maintain appealing moral identities in front of our multiple constituencies. Rationalists such as Kohlberg and Turiel portrayed children as little scientists who use logic and experimentation to figure out the truth for themselves. When we look at children's efforts to understand the physical world, the scientist metaphor is apt; kids really are formulating and testing hypotheses, and they really do converge, gradually, on the truth.<sup>9</sup> But in the social world, things are different, according to Tetlock. The social world is Glauconian.<sup>10</sup> Appearance is usually far more important than reality.

In Tetlock's research, subjects are asked to solve problems and make decisions.<sup>11</sup> For example, they're given information about a legal case and then asked to infer guilt or innocence. Some subjects are told that they'll have to explain their decisions to someone else. Other subjects know that they won't be held accountable by anyone. Tetlock found that when left to their own devices, people show the usual catalogue of errors, laziness, and reliance on gut feelings that has been documented in so much decision-making research.<sup>12</sup> But when people know in advance that they'll have to explain themselves, they think more systematically and self-critically. They are less likely

to jump to premature conclusions and more likely to revise their beliefs in response to evidence.

That might be good news for rationalists—maybe we can think carefully whenever we believe it matters? Not quite. Tetlock found two very different kinds of careful reasoning. *Exploratory thought* is an “evenhanded consideration of alternative points of view.” *Confirmatory thought* is “a one-sided attempt to rationalize a particular point of view.”<sup>13</sup> Accountability increases exploratory thought only when three conditions apply: (1) decision makers learn before forming any opinion that they will be accountable to an audience, (2) the audience’s views are unknown, and (3) they believe the audience is well informed and interested in accuracy.

When all three conditions apply, people do their damndest to figure out the truth, because that’s what the audience wants to hear. But the rest of the time—which is almost all of the time—accountability pressures simply increase confirmatory thought. People are trying harder to *look* right than to *be* right. Tetlock summarizes it like this:

A central function of thought is making sure that one acts in ways *that can be persuasively justified or excused to others*. Indeed, the process of considering the justifiability of one’s choices may be so prevalent that decision makers not only search for convincing reasons to make a choice when they must explain that choice to others, *they search for reasons to convince themselves* that they have made the “right” choice.<sup>14</sup>

Tetlock concludes that conscious reasoning is carried out largely for the purpose of persuasion, rather than discovery. But Tetlock adds that we are also trying to persuade ourselves. We want to believe the things we are about to say to others. In the rest of this chapter I’ll review five bodies of experimental research supporting Tetlock and Glaucon. Our moral thinking is much more like a politician searching for votes than a scientist searching for truth.

#### 1. WE ARE OBSESSED WITH POLLS

Ed Koch, the brash mayor of New York City in the 1980s, was famous for greeting constituents with the question “How’m I doin’?” It was a humorous reversal of the usual New York “How you doin’?” but it conveyed the chronic concern of elected officials. Few of us will ever run for office, yet most of the people we meet belong to one or more constituencies that we want to win over. Research on self-esteem suggests that we are all unconsciously asking Koch’s question every day, in almost every encounter.

For a hundred years, psychologists have written about the need to think well of oneself. But Mark Leary, a leading researcher on self-consciousness, thought that it made no evolutionary sense for there to be a deep need for *self-esteem*.<sup>15</sup> For millions of years, our ancestors’ survival depended upon their ability to get small groups to include them and trust them, so if there is any innate drive here, it should be a drive to get *others* to think well of us. Based on his review of the research, Leary suggested that self-esteem is more like an internal gauge, a “sociometer” that continuously measures your value as a relationship partner. Whenever the sociometer needle drops, it triggers an alarm and changes our behavior.

As Leary was developing the sociometer theory in the 1990s, he kept meeting people who denied that they were affected by what others thought of them. Do some people truly steer by their own compass?

Leary decided to put these self-proclaimed mavericks to the test. First, he had a large group of students rate their self-esteem and how much it depended on what other people think. Then he picked out the few people who—question after question—said they were completely unaffected by the opinions of others, and he invited them to the lab a few weeks later. As a comparison, he also invited people who had consistently said that they *were* strongly affected by what other people think of them. The test was on.

Everyone had to sit alone in a room and talk about themselves for five minutes, speaking into a microphone. At the end of each minute

they saw a number flash on a screen in front of them. That number indicated how much another person listening in from another room wanted to interact with them in the next part of the study. With ratings from 1 to 7 (where 7 is best), you can imagine how it would feel to see the numbers drop while you're talking: 4...3...2...3...3...2.

In truth, Leary had rigged it. He gave some people declining ratings while other people got rising ratings: 4...5...6...5...6. Obviously it's more pleasant to see your numbers rise, but would seeing either set of numbers (ostensibly from a complete stranger) change what you believe to be true about yourself, your merits, your self-worth?

Not surprisingly, people who admitted that they cared about other people's opinions had big reactions to the numbers. Their self-esteem sank. But the self-proclaimed mavericks suffered shocks almost as big: They might indeed have steered by their own compass, but they didn't realize that their compass tracked public opinion, not true north. It was just as Glaucon said.

Leary's conclusion was that "the sociometer operates at a nonconscious and preattentive level to scan the social environment for any and all indications that one's relational value is low or declining."<sup>6</sup> The sociometer is part of the elephant. Because appearing concerned about other people's opinions makes us look weak, we (like politicians) often deny that we care about public opinion polls. But the fact is that we care a lot about what others think of us. The only people known to have no sociometer are psychopaths.<sup>7</sup>

## 2. OUR IN-HOUSE PRESS SECRETARY AUTOMATICALLY JUSTIFIES EVERYTHING

If you want to see post hoc reasoning in action, just watch the press secretary of a president or prime minister take questions from reporters. No matter how bad the policy, the secretary will find some way to praise or defend it. Reporters then challenge assertions and bring up contradictory quotes from the politician, or even quotes straight from the press secretary on previous days. Sometimes you'll hear

an awkward pause as the secretary searches for the right words, but what you'll never hear is: "Hey, that's a great point! Maybe we should rethink this policy."

Press secretaries can't say that because they have no power to make or revise policy. They're told what the policy is, and their job is to find evidence and arguments that will justify the policy to the public. And that's one of the rider's main jobs: to be the full-time in-house press secretary for the elephant.

In 1960, Peter Wason (creator of the 4-card task from chapter 2) published his report on the "2-4-6 problem."<sup>8</sup> He showed people a series of three numbers and told them that the triplet conforms to a rule. They had to guess the rule by generating other triplets and then asking the experimenter whether the new triplet conformed to the rule. When they were confident they had guessed the rule, they were supposed to tell the experimenter their guess.

Suppose a subject first sees 2-4-6. The subject then generates a triplet in response: "4-6-8?"

"Yes," says the experimenter.

"How about 120-122-124?"

"Yes."

It seemed obvious to most people that the rule was consecutive even numbers. But the experimenter told them this was wrong, so they tested out other rules: "3-5-7?"

"Yes."

"What about 35-37-39?"

"Yes."

"OK, so the rule must be any series of numbers that rises by two?"

"No."

People had little trouble generating new hypotheses about the rule, sometimes quite complex ones. But what they hardly ever did was to test their hypotheses by offering triplets that *did not conform to their hypothesis*. For example, proposing 2-4-5 (yes) and 2-4-3 (no) would have helped people zero in on the actual rule: any series of ascending numbers.

Wason called this phenomenon the *confirmation bias*, the tendency to seek out and interpret new evidence in ways that confirm what you

already think. People are quite good at challenging statements made by *other* people, but if it's *your* belief, then it's your possession—your child, almost—and you want to protect it, not challenge it and risk losing it.<sup>19</sup>

Deanna Kuhn, a leading researcher of everyday reasoning, found evidence of the confirmation bias even when people solve a problem that is important for survival: knowing what foods make us sick. To bring this question into the lab she created sets of eight index cards, each of which showed a cartoon image of a child eating something—chocolate cake versus carrot cake, for example—and then showed what happened to the child afterward: the child is smiling, or else is frowning and looking sick. She showed the cards one at a time, to children and to adults, and asked them to say whether the “evidence” (the 8 cards) suggested that either kind of food makes kids sick.

The kids as well as the adults usually started off with a hunch—in this case, that chocolate cake is the more likely culprit. They usually concluded that the evidence proved them right. Even when the cards showed a stronger association between carrot cake and sickness, people still pointed to the one or two cards with sick chocolate cake eaters as evidence for their theory, and they ignored the larger number of cards that incriminated carrot cake. As Kuhn puts it, people seemed to say to themselves: “Here is some evidence I can point to as supporting my theory, and therefore the theory is right.”<sup>20</sup>

This is the sort of bad thinking that a good education should correct, right? Well, consider the findings of another eminent reasoning researcher, David Perkins.<sup>21</sup> Perkins brought people of various ages and education levels into the lab and asked them to think about social issues, such as whether giving schools more money would improve the quality of teaching and learning. He first asked subjects to write down their initial judgment. Then he asked them to think about the issue and write down all the reasons they could think of—on either side—that were relevant to reaching a final answer. After they were done, Perkins scored each reason subjects wrote as either a “my-side” argument or an “other-side” argument.

Not surprisingly, people came up with many more “my-side” arguments than “other-side” arguments. Also not surprisingly, the

more education subjects had, the more reasons they came up with. But when Perkins compared fourth-year students in high school, college, or graduate school to first-year students in those same schools, he found barely any improvement within each school. Rather, the high school students who generate a lot of arguments are the ones who are more likely to go on to college, and the college students who generate a lot of arguments are the ones who are more likely to go on to graduate school. Schools don't *teach* people to reason thoroughly; they *select* the applicants with higher IQs, and people with higher IQs are able to generate more reasons.

The findings get more disturbing. Perkins found that IQ was by far the biggest predictor of how well people argued, but it predicted *only the number of my-side arguments*. Smart people make really good lawyers and press secretaries, but they are no better than others at finding reasons on the other side. Perkins concluded that “people invest their IQ in buttressing their own case rather than in exploring the entire issue more fully and evenhandedly.”<sup>22</sup>

Research on everyday reasoning offers little hope for moral rationalists. In the studies I've described, there is no self-interest at stake. When you ask people about strings of digits, cakes and illnesses, and school funding, people have rapid, automatic intuitive reactions. One side looks a bit more attractive than the other. The elephant leans, ever so slightly, and the rider gets right to work looking for supporting evidence—and invariably succeeds.

This is how the press secretary works on trivial issues where there is no motivation to support one side or the other. If thinking is confirmatory rather than exploratory in these dry and easy cases, then what chance is there that people will think in an open-minded, exploratory way when self-interest, social identity, and strong emotions make them want or even *need* to reach a preordained conclusion?

### 3. WE LIE, CHEAT, AND JUSTIFY SO WELL THAT WE HONESTLY BELIEVE WE ARE HONEST

In the United Kingdom, members of Parliament (MPs) have long been allowed to bill taxpayers for the reasonable expense of maintaining a second home, given that they're required to spend time in London and in their home districts. But because the office responsible for deciding what was reasonable approved nearly every request, members of Parliament treated it like a big blank check. And because their expenses were hidden from the public, MPs thought they were wearing the ring of Gyges—until a newspaper printed a leaked copy of those expense claims in 2009.<sup>23</sup>

Just as Glaucon predicted, they had behaved abominably. Many MPs declared their second home to be whichever one was due for major and lavish renovations (including dredging the moats). When the renovations were completed, they simply redesignated their primary home as their secondary home and renovated that one too, sometimes selling the newly renovated home for a huge profit.

Late-night comedians are grateful for the never-ending stream of scandals coming out of London, Washington, and other centers of power. But are the rest of us any better than our leaders? Or should we first look for logs in our own eyes?

Many psychologists have studied the effects of having “plausible deniability.” In one such study, subjects performed a task and were then given a slip of paper and a verbal confirmation of how much they were to be paid. But when they took the slip to another room to get their money, the cashier mistread one digit and handed them too much money. Only 20 percent spoke up and corrected the mistake.<sup>24</sup>

But the story changed when the cashier asked them if the payment was correct. In that case, 60 percent said no and returned the extra money. Being asked directly removes plausible deniability; it would take a direct lie to keep the money. As a result, people are three times more likely to be honest.

You can't predict who will return the money based on how people rate their own honesty, or how well they are able to give the

high-minded answer on a moral dilemma of the sort used by Kohlberg.<sup>25</sup> If the rider were in charge of ethical behavior, then there would be a big correlation between people's moral reasoning and their moral behavior. But he's not, so there isn't.

In his book *Predictably Irrational*, Dan Ariely describes a brilliant series of studies in which participants had the opportunity to earn more money by claiming to have solved more math problems than they really did. Ariely summarizes his findings from many variations of the paradigm like this:

When given the opportunity, many honest people will cheat. In fact, rather than finding that a few bad apples weighted the averages, we discovered that *the majority of people cheated*, and that they cheated just a little bit.<sup>26</sup>

People didn't try to get away with as much as they could. Rather, when Ariely gave them anything like the invisibility of the ring of Gyges, they cheated only up to the point where they themselves could no longer find a justification that would preserve their belief in their own honesty.

The bottom line is that in lab experiments that give people invisibility combined with plausible deniability, *most people cheat*. The press secretary (also known as the *inner lawyer*)<sup>27</sup> is so good at finding justifications that most of these cheaters leave the experiment as convinced of their own virtue as they were when they walked in.

### 4. REASONING (AND GOOGLE) CAN TAKE YOU WHEREVER YOU WANT TO GO

When my son, Max, was three years old, I discovered that he's allergic to *must*. When I would tell him that he *must* get dressed so that we can go to school (and he loved to go to school), he'd scowl and whine. The word *must* is a little verbal handcuff that triggered in him the desire to squirm free.

The word *can* is so much nicer: “Can you get dressed, so that



we can go to school?" To be certain that these two words were really right and day, I tried a little experiment. After dinner one night, I said "Max, you *must* eat ice cream now."

"But I don't want to!"

Four seconds later: "Max, you can have ice cream if you want."

"I want some!"

The difference between *can* and *must* is the key to understanding the profound effects of self-interest on reasoning. It's also the key to understanding many of the strangest beliefs—in UFO abductions, quack medical treatments, and conspiracy theories.

The social psychologist Tom Gilovich studies the cognitive mechanisms of strange beliefs. His simple formulation is that when we *want* to believe something, we ask ourselves, "*Can* I believe it?"<sup>28</sup> Then (as Kuhn and Perkins found), we search for supporting evidence, and if we find even a single piece of pseudo-evidence, we can stop thinking. We now have permission to believe. We have a justification, in case anyone asks.

In contrast, when we *don't* want to believe something, we ask ourselves, "*Must* I believe it?" Then we search for contrary evidence, and if we find a single reason to doubt the claim, we can dismiss it. You only need one key to unlock the handcuffs of *must*.

Psychologists now have file cabinets full of findings on "motivated reasoning,"<sup>29</sup> showing the many tricks people use to reach the conclusions they want to reach. When subjects are told that an intelligence test gave them a low score, they choose to read articles criticizing (rather than supporting) the validity of IQ tests.<sup>30</sup> When people read a (fictitious) scientific study that reports a link between caffeine consumption and breast cancer, women who are heavy coffee drinkers find more flaws in the study than do men and less caffeinated women.<sup>31</sup> Pete Ditto, at the University of California at Irvine, asked subjects to lick a strip of paper to determine whether they have a serious enzyme deficiency. He found that people wait longer for the paper to change color (which it never does) when a color change is desirable than when it indicates a deficiency, and those who get the undesirable prognosis find more reasons why the test might not be accurate (for example, "My mouth was unusually dry today").<sup>32</sup>

The difference between a mind asking "Must I believe it?" versus "Can I believe it?" is so profound that it even influences visual perception. Subjects who thought that they'd get something good if a computer flashed up a letter rather than a number were more likely to see the ambiguous figure **B** as the letter *B*, rather than as the number 13.<sup>33</sup>

If people can literally see what they want to see—given a bit of ambiguity—is it any wonder that scientific studies often fail to persuade the general public? Scientists are really good at finding flaws in studies that contradict their own views, but it sometimes happens that evidence accumulates across many studies to the point where scientists *must* change their minds. I've seen this happen in my colleagues (and myself) many times,<sup>34</sup> and it's part of the accountability system of science—you'd look foolish clinging to discredited theories. But for nonscientists, there is no such thing as a study you must believe. It's *always* possible to question the methods, find an alternative interpretation of the data, or, if all else fails, question the honesty or ideology of the researchers.

And now that we all have access to search engines on our cell phones, we can call up a team of supportive scientists for almost any conclusion twenty-four hours a day. Whatever you want to believe about the causes of global warming or whether a fetus can feel pain, just Google your belief. You'll find partisan websites summarizing and sometimes distorting relevant scientific studies. Science is a smorgasbord, and Google will guide you to the study that's right for you.

##### 5. WE CAN BELIEVE ALMOST ANYTHING THAT SUPPORTS OUR TEAM

Many political scientists used to assume that people vote selfishly, choosing the candidate or policy that will benefit them the most. But decades of research on public opinion have led to the conclusion that self-interest is a weak predictor of policy preferences. Parents of children in public school are not more supportive of government aid to schools than other citizens; young men subject to the draft

are not more opposed to military escalation than men too old to be drafted; and people who lack health insurance are not more likely to support government-issued health insurance than people covered by insurance.<sup>35</sup>

Rather, people care about their *groups*, whether those be racial, regional, religious, or political. The political scientist Don Kinder summarizes the findings like this: "In matters of public opinion, citizens seem to be asking themselves not 'What's in it for me?' but rather 'What's in it for my group?'"<sup>36</sup> Political opinions function as "badges of social membership."<sup>37</sup> They're like the array of bumper stickers people put on their cars showing the political causes, universities, and sports teams they support. Our politics is groupish, not selfish.

If people can see what they want to see in the figure **B**, just imagine how much room there is for partisans to see different facts in the social world.<sup>38</sup> Several studies have documented the "attitude polarization" effect that happens when you give a single body of information to people with differing partisan leanings. Liberals and conservatives actually move further apart when they read about research on whether the death penalty deters crime, or when they rate the quality of arguments made by candidates in a presidential debate, or when they evaluate arguments about affirmative action or gun control.<sup>39</sup>

In 2004, in the heat of the U.S. presidential election, Drew Westen used fMRI to catch partisan brains in action.<sup>40</sup> He recruited fifteen highly partisan Democrats and fifteen highly partisan Republicans and brought them into the scanner one at a time to watch eighteen sets of slides. The first slide in each set showed either a statement from President George W. Bush or one from his Democratic challenger, John Kerry. For example, people saw a quote from Bush in 2000 praising Ken Lay, the CEO of Enron, which later collapsed when its massive frauds came to light:

I love the man. . . . When I'm president, I plan to run the government like a CEO runs a country. Ken Lay and Enron are a model of how I'll do that.

Then they saw a slide describing an action taken later that seemed to contradict the earlier statement:

Mr. Bush now avoids any mention of Ken Lay, and is critical of Enron when asked.

At this point, Republicans were squirming. But right then, Westen showed them another slide that gave more context, resolving the contradiction:

People who know the President report that he feels betrayed by Ken Lay, and was genuinely shocked to find that Enron's leadership had been corrupt.

There was an equivalent set of slides showing Kerry caught in a contradiction and then released. In other words, Westen engineered situations in which partisans would temporarily feel threatened by their candidates' apparent hypocrisy. At the same time, they'd feel no threat—and perhaps even pleasure—when it was the other party's guy who seemed to have been caught.

Westen was actually pitting two models of the mind against each other: Would subjects reveal Jefferson's dual-process model, in which the head (the reasoning parts of the brain) processes information about contradictions equally for all targets, but then gets overruled by a stronger response from the heart (the emotion areas)? Or does the partisan brain work as Hume says, with emotional and intuitive processes running the show and only putting in a call to reasoning when its services are needed to justify a desired conclusion?

The data came out strongly supporting Hume. The threatening information (their own candidate's hypocrisy) immediately activated a network of emotion-related brain areas—areas associated with negative emotion and responses to punishment.<sup>41</sup> The handcuffs (of "Must I believe it?") hurt.

Some of these areas are known to play a role in reasoning, but there was no increase in activity in the dorso-lateral prefrontal cortex (dlPFC). The dlPFC is the main area for cool reasoning tasks.<sup>42</sup>

Whatever thinking partisans were doing, it was not the kind of objective weighing or calculating that the dlPFC is known for.<sup>43</sup>

Once Westen released them from the threat, the ventral striatum started humming—that's one of the brain's major reward centers. All animal brains are designed to create flashes of pleasure when the animal does something important for its survival, and small pulses of the neurotransmitter dopamine in the ventral striatum (and a few other places) are where these good feelings are manufactured. Heroin and cocaine are addictive because they artificially trigger this dopamine response. Rats who can press a button to deliver electrical stimulation to their reward centers will continue pressing until they collapse from starvation.<sup>44</sup>

Westen found that partisans escaping from handcuffs (by thinking about the final slide, which restored their confidence in their candidate) got a little hit of that dopamine. And if this is true, then it would explain why extreme partisans are so stubborn, closed-minded, and committed to beliefs that often seem bizarre or paranoid. Like rats that cannot stop pressing a button, partisans may be simply unable to stop believing weird things. The partisan brain has been reinforced so many times for performing mental contortions that free it from unwanted beliefs. Extreme partisanship may be literally addictive.

#### THE RATIONALIST DELUSION

*Webster's Third New International Dictionary* defines *delusion* as "a false conception and persistent belief unconquerable by reason in something that has no existence in fact."<sup>45</sup> As an intuitionist, I'd say that the worship of reason is itself an illustration of one of the most long-lived delusions in Western history: the rationalist delusion. It's the idea that reasoning is our most noble attribute, one that makes us like the gods (for Plato) or that brings us beyond the "delusion" of believing in gods (for the New Atheists).<sup>46</sup> The rationalist delusion is not just a claim about human nature. It's also a claim that the rational caste (philosophers or scientists) should have more power, and it usually comes along with a utopian program for raising more rational children.<sup>47</sup>

From Plato through Kant and Kohlberg, many rationalists have asserted that the ability to reason well about ethical issues *causes* good behavior. They believe that reasoning is the royal road to moral truth, and they believe that people who reason well are more likely to act morally.

But if that were the case, then moral philosophers—who reason about ethical principles all day long—should be more virtuous than other people. Are they? The philosopher Eric Schwitzgebel tried to find out. He used surveys and more surreptitious methods to measure how often moral philosophers give to charity, vote, call their mothers, donate blood, donate organs, clean up after themselves at philosophy conferences, and respond to emails purportedly from students.<sup>48</sup> And in none of these ways are moral philosophers better than other philosophers or professors in other fields.

Schwitzgebel even scrounged up the missing-book lists from dozens of libraries and found that academic books on ethics, which are presumably borrowed mostly by ethicists, are more likely to be stolen or just never returned than books in other areas of philosophy.<sup>49</sup> In other words, expertise in moral reasoning does not seem to improve moral behavior, and it might even make it worse (perhaps by making the rider more skilled at post hoc justification). Schwitzgebel still has yet to find a single measure on which moral philosophers behave better than other philosophers.

Anyone who values truth should stop worshipping reason. We all need to take a cold hard look at the evidence and see reasoning for what it is. The French cognitive scientists Hugo Mercier and Dan Sperber recently reviewed the vast research literature on motivated reasoning (in social psychology) and on the biases and errors of reasoning (in cognitive psychology). They concluded that most of the bizarre and depressing research findings make perfect sense once you see reasoning as having evolved not to help us find truth but to help us engage in arguments, persuasion, and manipulation in the context of discussions with other people. As they put it, "skilled arguers . . . are not after the truth but after arguments supporting their views."<sup>50</sup> This explains why the confirmation bias is so powerful, and so ineradicable. How hard could it be to teach students to look on the other side, to look for evidence against their favored view? Yet, in

fact, it's very hard, and nobody has yet found a way to do it.<sup>51</sup> It's hard because the confirmation bias is a built-in feature (of an argumentative mind), not a bug that can be removed (from a platonic mind).

I'm not saying we should all stop reasoning and go with our gut feelings. Gut feelings are sometimes better guides than reasoning for making consumer choices and interpersonal judgments,<sup>52</sup> but they are often disastrous as a basis for public policy, science, and law.<sup>53</sup> Rather, what I'm saying is that we must be wary of any *individual's* ability to reason. We should see each individual as being limited, like a neuron. A neuron is really good at one thing: summing up the stimulation coming into its dendrites to "decide" whether to fire a pulse along its axon. A neuron by itself isn't very smart. But if you put neurons together in the right way you get a brain; you get an emergent system that is much smarter and more flexible than a single neuron.

In the same way, each individual reasoner is really good at one thing: finding evidence to support the position he or she already holds, usually for intuitive reasons. We should not expect individuals to produce good, open-minded, truth-seeking reasoning, particularly when self-interest or reputational concerns are in play. But if you put individuals together in the right way, such that some individuals can use their reasoning powers to disconfirm the claims of others, and all individuals feel some common bond or shared fate that allows them to interact civilly, you can create a group that ends up producing good reasoning as an emergent property of the social system. This is why it's so important to have intellectual and ideological diversity within any group or institution whose goal is to find truth (such as an intelligence agency or a community of scientists) or to produce good public policy (such as a legislature or advisory board).

And if our goal is to produce good *behavior*, not just good thinking, then it's even more important to reject rationalism and embrace intuitionism. Nobody is ever going to invent an ethics class that makes people behave ethically after they step out of the classroom. Classes are for riders, and riders are just going to use their new knowledge to serve their elephants more effectively. If you want to make people behave more ethically, there are two ways you can go. You can change the elephant, which takes a long time and is hard to do. Or, to borrow an idea from the book *Swirls*, by Chip Heath and Dan Heath,<sup>54</sup> you

can change the path that the elephant and rider find themselves traveling on. You can make minor and inexpensive tweaks to the environment, which can produce big increases in ethical behavior.<sup>55</sup> You can hire Glaucon as a consultant and ask him how to design institutions in which real human beings, always concerned about their reputations, will behave more ethically.

#### IN SUM

The first principle of moral psychology is *Intuitions come first, strategic reasoning second*. To demonstrate the strategic functions of moral reasoning, I reviewed five areas of research showing that moral thinking is more like a politician searching for votes than a scientist searching for truth:

- We are obsessively concerned about what others think of us, although much of the concern is unconscious and invisible to us.
- Conscious reasoning functions like a press secretary who automatically justifies any position taken by the president.
- With the help of our press secretary, we are able to lie and cheat often, and then cover it up so effectively that we convince even ourselves.
- Reasoning can take us to almost any conclusion we want to reach, because we ask "Can I believe it?" when we want to believe something, but "Must I believe it?" when we don't want to believe. The answer is almost always yes to the first question and no to the second.
- In moral and political matters we are often groupish, rather than selfish. We deploy our reasoning skills to support our team, and to demonstrate commitment to our team.

I concluded by warning that the worship of reason, which is sometimes found in philosophical and scientific circles, is a delusion.

It is an example of faith in something that does not exist. I urged instead a more intuitionist approach to morality and moral education, one that is more humble about the abilities of individuals, and more attuned to the contexts and social systems that enable people to think and act well.

I have tried to make a reasoned case that our moral capacities are best described from an intuitionist perspective. I do not claim to have examined the question from all sides, nor to have offered irrefutable proof. Because of the insurmountable power of the confirmation bias, counterarguments will have to be produced by those who disagree with me. Eventually, if the scientific community works as it is supposed to, the truth will emerge as a large number of flawed and limited minds battle it out.

This concludes Part I of this book, which was about the first principle of moral psychology: *Intuitions come first, strategic reasoning second*. To explain this principle I used the metaphor of the mind as a rider (reasoning) on an elephant (intuition), and I said that the rider's function is to serve the elephant. Reasoning matters, particularly because reasons do sometimes influence other people, but most of the action in moral psychology is in the intuitions. In Part II I'll get much more specific about what those intuitions are and where they came from. I'll draw a map of moral space, and I'll show why that map is usually more favorable to conservative politicians than to liberals.

## PART II

# There's More to Morality than Harm and Fairness

## Central Metaphor

*The righteous mind is like a tongue  
with six taste receptors.*

28. Hare 1993.
29. *Ibid.*, p. 54.
30. *Ibid.*, p. 91.
31. Beaver et al. 2011; Blomigen et al. 2005; Viding et al. 2005.
32. Brain scanning studies confirm that many emotional areas, including the amygdala and the vmPFC, are much less reactive in psychopaths than in normal people; see Blair 2007; Kiehl 2006. If you hook them up to a skin conductance meter, as in a lie detector test, psychopaths show a normal response to a photograph of a shark with open jaws. But show them a picture of mutilated bodies or suffering children, and the meter doesn't budge (Blair 1999). For the best clinical portraits of psychopaths and their indifference to others, including their parents, see Cleckley 1955.
33. James 1950/1890, I:488.
34. Baillargeon 1987.
35. The first work demonstrating that infants have innate abilities to understand the social world, including abilities to infer intentions and react to harm, was done by David and Ann Premack; see Premack and Premack 1994 for a review summarizing the origins of moral cognition.
36. Hamlin, Wynn, and Bloom 2007. This looking-time difference was found only for the ten-month-old children, not the six-month-olds. But the reaching-out difference was found for both age groups. The puppets were not traditional puppets; they were different colors and shapes of wood blocks. You can view the puppet shows from links at [www.yale.edu/infantlab/In\\_the\\_Media.html](http://www.yale.edu/infantlab/In_the_Media.html). This technique of measuring infants' attributions was first developed by Kuhlmeier, Wynn, and Bloom 2003.
37. Hamlin, Wynn, and Bloom 2007, p. 559.
38. For early writings on this idea, see Hoffman 1982; Kagan 1984.
39. The trolley dilemma was first discussed by philosophers Philippa Foot and Judith Jarvis Thompson.
40. Some philosophers note the difference that in the bridge story you are using the victim as a means to an end, whereas in the switch story the victim is not a means to an end; his death is just an unfortunate side effect. Greene and others have therefore tested alternative versions, such as the case where the switch only saves lives because it diverts the trolley onto a side loop where one man is standing. In that case the victim is still being used as a means to an end; if he were to step off the track, the trolley would continue on the loop, back onto the main track, and would kill the five people. In these cases, subjects tend to give responses in between the original switch and footbridge versions.
41. Greene et al. 2001. This study also reported that it took longer for subjects who did make the utilitarian choice to give their answer, as though reasoning was

- struggling to overcome emotion, although that finding was later shown to be an artifact of the particular stories chosen, not a general principle (McGuire et al. 2009). But see Greene 2009b for a response.
42. Rilling et al. 2008; Sanfey et al. 2003.
  43. For reviews see Greene 2009a and Greene forthcoming. The areas most frequently reported include the vmPFC, insula, and amygdala. For an exception, see Knoch, Pascual-Leone, Meyer, Treyer, and Fehr 2006.
  44. Greene 2008; the quote is on p. 63. I asked Greene if he had known about the Wilson quote from p. 563 of *Sociobiology*, and he said no.
  45. See my review of these works in Haidt and Kesebir 2010.
  46. See Sinnott-Armstrong 2008 for a three-volume set of papers by this interdisciplinary community.
  47. Paxton, Ungar, and Greene, forthcoming.
  48. I should note that people vary in the degree to which they feel strong intuitions, in their ability to construct reasons, and in their openness to the reasons of others. See Bartels 2008 for a discussion of these individual differences.
4. VOTE FOR ME (HERE'S WHY)
1. *Republic*, 360c, trans. G. M. A. Grube and C. D. C. Reeve. In Plato 1997.
  2. It is Glaucon's brother Adeimantus who states the challenge in this way, at 360e–361d, but he's just elaborating upon Glaucon's argument. Glaucon and Adeimantus want Socrates to succeed and refute their arguments. Nonetheless, I will use Glaucon for the rest of this book as a spokesman for the view that reputation matters more than reality.
  3. *Republic*, 443–45.
  4. *Ibid.*, 473ff.
  5. At least Plato stated his assumptions about human nature at great length. Many other moral philosophers, such as Kant and Rawls, simply make assertions about how minds work, what people want, or what seems "reasonable." These assertions seem to be based on little more than introspection about their own rather unusual personalities or value systems. For example, when some of Rawls's (1971) assumptions were tested—e.g., that most people would care more about raising the worst-off than about raising the average if they had to design a society from behind a "veil of ignorance," so that they don't know what position they'd occupy in the society—they were found to be false (Frohlich, Oppenheimer, and Eavey 1987).
  6. His exact words were: "My thinking is first and last and always for the sake of my doing" (James 1950/1890, p. 333). Susan Fiske (1993) applied James's functionalism to social cognition, abbreviating his dictum as "thinking is for doing." For more on functionalism in the social sciences, see Merton 1968.

7. A rationalist can still believe that reasoning is easily corrupted, or that most people don't reason properly. But ought implies can, and rationalists are committed to the belief that reason *can* work this way, perhaps (as in Plato's case) because perfect rationality is the soul's true nature.
8. Lerner and Tetlock 2003, p. 434.
9. Gopnik, Meltzoff, and Kuhl 2000.
10. I could perhaps use the term *Machiavellian* instead of *Glauconian* throughout this book. But the word *Machiavellian* is too dark, too suggestive of leaders tricking people in order to dominate them. I think moral life is really about cooperation and alliance, rather than about power and domination. The dishonesty and hypocrisy of our moral reasoning is done to get people to like us and cooperate with us, so I prefer the term *Glauconian*.
11. See review in Lerner and Tetlock 2003. Tetlock 2002 presents three metaphors: intuitive politicians, intuitive prosecutors, and intuitive theologians. I focus on the intuitive politician here, and I present the intuitive prosecutor below, as being related to the needs of the intuitive politician. I cover the subject matter of the intuitive theologian when I discuss religion and the need to bind people together with shared beliefs about sacredness, in chapter 11.
12. For reviews see Ariely 2008; Baron 2007.
13. Lerner and Tetlock 2003, p. 438.
14. *Ibid.*, p. 433; emphasis added.
15. Leary 2004.
16. Leary 2005, p. 85. There surely are differences among people in how obsessed they are with the opinions of others. But Leary's findings indicate that we are not particularly accurate at assessing our own levels of obsession.
17. Milion et al. 1998. Psychopaths often care what others think, but only as part of a plan to manipulate or exploit others. They don't have emotions such as shame and guilt that make it painful for them when others see through their lies and come to hate them. They don't have an automatic unconscious sociometer.
18. Wason 1960.
19. Shaw 1996. The confirmation bias is found widely in social, clinical, and cognitive psychology. It appears early in childhood and it lasts for life. See reviews in Kunda 1990; Mercier & Sperber 2010; Nickerson 1998; Pyszczynski and Greenberg 1987.
20. Kahn 1989, p. 681.
21. Perkins, Farady, and Bushey 1991.
22. *Ibid.*, p. 95. They did find a bit of overall improvement between the first and fourth year of high school, but this might have been simple maturation, rather than an effect of education. They didn't find it in college.
23. The *Daily Telegraph* got a leaked copy of the full expense report, which had

- been prepared by the House of Commons in response to a Freedom of Information request that it had resisted for years.
24. Bersoff 1999. See also Dan Batson's research on "moral hypocrisy," e.g., Batson et al. 1999.
  25. Pergini and Leone 2009.
  26. Ariely 2008, p. 201; emphasis added.
  27. This is the term I used in *The Happiness Hypothesis*.
  28. Gilovich 1991, p. 84.
  29. Ditto, Pizarro, and Tannenbaum 2009; Kunda 1990.
  30. Frey and Stahberg 1986.
  31. Kunda 1987.
  32. Ditto and Lopez 1992. See also Ditto et al. 2003, which finds that when we want to believe something, we often don't even bother to search for a single piece of supporting evidence. We just accept things uncritically.
  33. Balcells and Dunning 2006.
  34. See Brockman 2009.
  35. See review in Kinder 1998. The exception to this rule is that when the material benefits of a policy are "substantial, imminent, and well-publicized," those who would benefit from it are more likely to support it than those who would be harmed. See also D. T. Miller 1999 on the "norm of self-interest."
  36. Kinder 1998, p. 808.
  37. The term is from Smith, Bruner, and White, as quoted by Kinder 1998.
  38. See the classic study by Hastorf and Cantril (1954) in which students at Dartmouth and Princeton came to very different conclusions about what had happened on the football field after watching the same film showing several disputed penalty calls.
  39. Lord, Ross, and Lepper 1979; Munro et al. 2002; Taber and Lodge 2006. Polarization effects are not found in all studies, but as Taber and Lodge argue, the studies that failed to find the effect generally used cooler, less emotional stimuli that did not fully engage partisan motivations.
  40. Westen et al. 2006.
  41. The activated areas included insula, medial PFC, ventral ACC, ventromedial PFC, and posterior cingulate cortex. The areas associated with negative emotion are particularly the left insula, lateral orbital frontal cortex, and ventromedial PFC. The amygdala, closely related to fear and threat, did show greater activity in the early trials but had "habitated" in the later trials. Note that all of these findings come from subtracting reactions to hypocrisy by the neutral target (e.g., Tom Hanks) from reactions to hypocrisy by one's own candidate.
  42. Greene (2008) refers to this area as "Mill" in the brain, because it tends to be

- more active when subjects make the cool, utilitarian choice, rather than the more emotion-based deontological choice.
43. The dIPFC did not show an increase in activity until *after* the eculpatory information was given and the partisan was freed from the handcuffs. It was as if confirmatory reasoning could not even begin until subjects had a clear and emotionally acceptable explanation to confirm.
  44. Olds and Milner 1954.
  45. *Weber's Third New International Dictionary*. Related definitions include “false belief or a persistent error of perception occasioned by a false belief or mental derangement.”
  46. Dawkins 2006; Dennett 2006; Harris 2006. I'll discuss their arguments in detail in chapter 11.
  47. Plato gives his childrearing advice in Book 3 of *The Republic*; Dawkins gives it in chapter 9 of *The God Delusion*.
  48. Schwitzgebel and Rust 2009, 2011; Schwitzgebel et al. 2011.
  49. Schwitzgebel 2009.
  50. Mercier and Sperber 2011, p. 57.
  51. See Lilienfeld, Ammirati, and Landfield 2009 for a report on how hard it has been to develop methods of “debiasing” human thinking. What little success there is in the “critical thinking” literature almost never finds (or even looks for) transfer of skills beyond the classroom.
  52. Wilson 2002; Wilson and Schooler 1991.
  53. Baron 1998.
  54. Heath and Heath 2010.
  55. See [www.EthicalSystems.org](http://www.EthicalSystems.org) for my attempt to bring together research on these “path changes,” many of which are simple to do. One good example is Dan Ariely's finding that if you ask people to sign an expense report at the beginning, promising to be honest, rather than at the end, affirming that they were honest, you get a big drop in overclaiming of expenses. See Ariely 2008.

#### 5. BEYOND WEIRD MORALITY

1. Mill 2003/1859, p. 80.
2. Henrich, Heine, and Norenzayan 2010.
3. Markus and Kitayama 1991.
4. For a review of these sorts of cultural differences, see Kitayama et al. 2009.
5. Nisbett et al. 2001.
6. In *Analects* 15:24, Confucius is asked whether there is a single word that could guide one's life. He responds: “Should it not be reciprocity? What you do not wish for yourself, do not do to others.” (*Lays* 1997). But there is no way to reduce the moral teachings of the *Analects* to the golden rule. As I read them,

- the *Analects* rely upon all six of the moral foundations I'll present in chapters 7 and 8.
7. See, for example, the books of Sam Harris, such as *The End of Faith* and *The Moral Landscape*.
  8. Not entirely new. As Shweder 1990a explains, it has arisen several times in psychology. But if someone today calls herself a cultural psychologist, she probably orients herself to the field as it was reborn in the ten years after the publication of Shweder and LeVine 1984.
  9. Shweder 1990a.
  10. The first published mention of the three ethics was Shweder 1990b. The major statement of the theory is Shweder et al. 1997.
  11. Peter Singer is the most prominent utilitarian philosopher of our time. See P. Singer 1979.
  12. It need not be a soul in anything like the Christian sense. As Paul Bloom (2004) has shown, we are “natural born dualists.” Despite wide religious variations, most people (including many atheists) believe that the mind, spirit, or soul is something separable from the body, something that inhabits the body.
  13. This, for example, was the conclusion drawn by Sayyid Qurb, an Egyptian who spent two years studying in America in the 1940s. He was repulsed, and this moral repulsion influenced his later work as an Islamist philosopher and theorist, one of the main inspirations for Osama bin Laden and Al-Qaeda.
  14. These text analyses are reported in Haidt et al. 1993. See also work by Lene Arnett Jensen (1997, 1998), which reached similar findings applying Shweder's three ethics to differences between progressive and orthodox participants, in India and in the United States.
  15. I am forever grateful to the late Sukumar Sen and his son Surojit Sen, of Cuttack and Bhubaneswar, for their generosity and kindness.
  16. In the Koran, see 2:222, 4:43, 24:30. In the Hebrew Bible, see the book of Leviticus in particular. For Christianity, see Thomas 1983, chapter 1. Also see New Testament passages on the purifications of Jesus and his followers, e.g., John 3:25, 11:55; Acts 15:9, 20:26, 21:26, 24:18.
  17. We also wanted to explain why so many languages extend their word for “disgust” to apply not just to physically repulsive things like excrement but also to some moral violations—but not all violations, and not always the same ones across cultures (Haidt et al. 1997).
  18. People intuitively associate up with good and down with bad, even when up and down are just relative positions on a computer monitor (Meier and Robinson 2004). For overviews of research on this psychological dimension, see Brandt and Reyna 2011, Rozin, Haidt, and McCauley 2008, and chapter 9 of *The Happiness Hypothesis*.