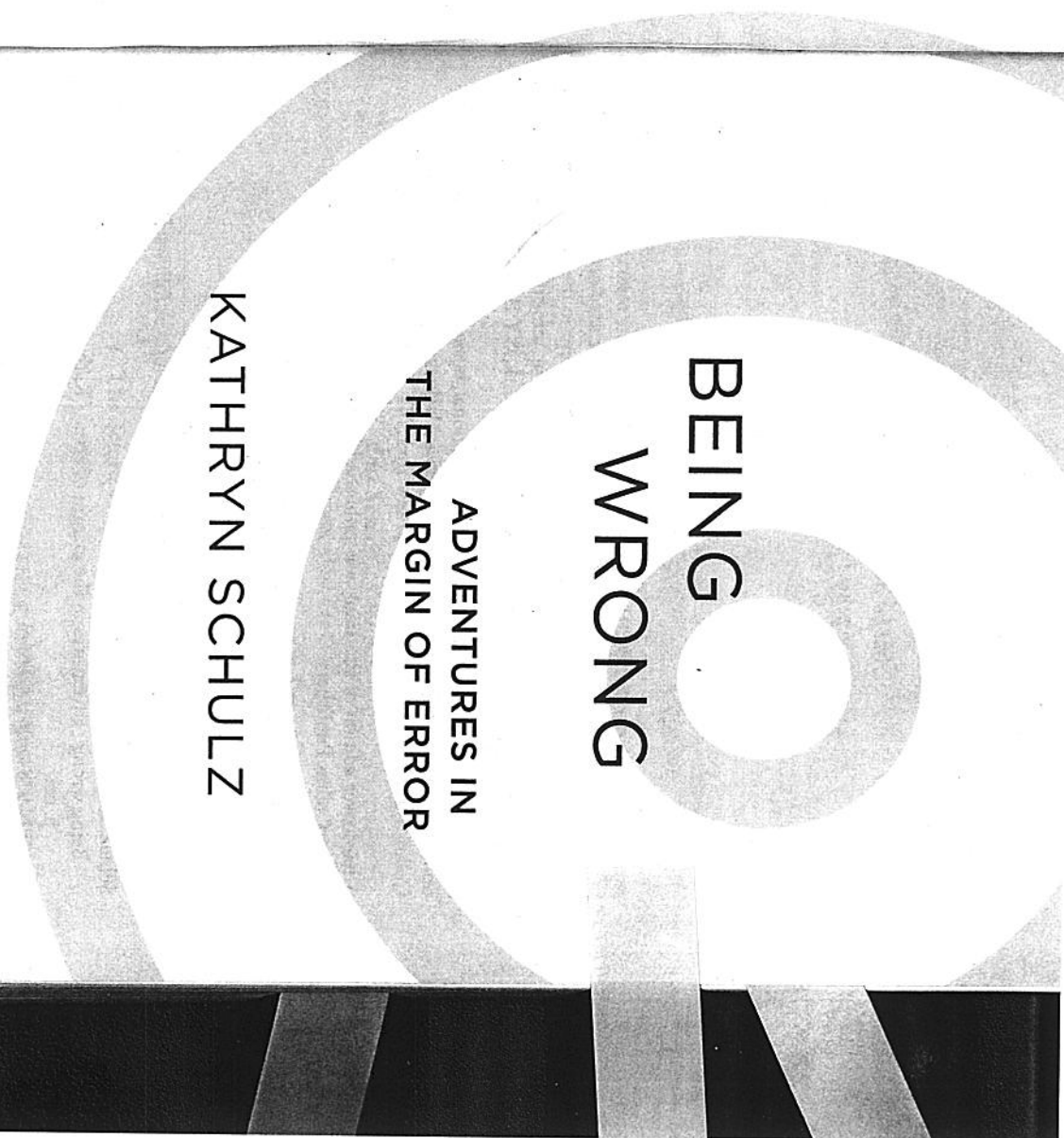


An abstract graphic design featuring a large, light gray spiral that originates from the center of the page and expands outwards. Overlaid on this spiral are several concentric circles of varying shades of gray. The text is positioned within the white space of these circles.

BEING WRONG

ADVENTURES IN
THE MARGIN OF ERROR

KATHRYN SCHULZ



An Imprint of HarperCollinsPublishers

*For my family, given and chosen
And for Michael and Amanda,
at whose expense I wrote about what I knew*

BEING WRONG. Copyright © 2010 by Kathryn Schulz. All rights reserved. Printed in the United States of America. No part of this book may be used or reproduced in any manner whatsoever without written permission except in the case of brief quotations embodied in critical articles and reviews. For information, address HarperCollins Publishers, 10 East 53rd Street, New York, NY 10022.

HarperCollins books may be purchased for educational, business, or sales promotional use. For information, please write: Special Markets Department, HarperCollins Publishers, 10 East 53rd Street, New York, NY 10022.

A hardcover edition of this book was published in 2010 by Ecco, an imprint of HarperCollins Publishers.

FIRST ECCO PAPERBACK EDITION PUBLISHED 2011.

Designed by Suet Yee Chong

The Library of Congress has cataloged the hardcover edition as follows:

Schulz, Kathryn.

Being wrong : adventures in the margin of error / Kathryn Schulz.—1st ed.
p. cm.

ISBN 978-0-06-117604-3

1. Fallibility. 2. Errors. 3. Decision making—Psychological aspects.
4. Philosophical anthropology. I. Title.

[call number n/a]
128 22

ISBN 978-0-06-117605-0 (pbk.) 2010282589

11 12 13 14 15 ID/RRD 10 9 8 7 6 5 4 3

CONTENTS

PART I THE IDEA OF ERROR

- 1 Wrongology 3
- 2 Two Models of Wrongness 25

PART II THE ORIGINS OF ERROR

- 3 Our Senses 47
- 4 Our Minds, Part One:
Knowing, Not Knowing, and Making It Up 67
- 5 Our Minds, Part Two: Belief 87
- 6 Our Minds, Part Three: Evidence 111
- 7 Our Society 133
- 8 The Allure of Certainty 159

PART III THE EXPERIENCE OF ERROR

- 9 Being Wrong 183
- 10 How Wrong? 201
- 11 Denial and Acceptance 220
- 12 Heartbreak 247
- 13 Transformation 273

easily into error. In other words, we have no sound method for knowing what we know—which means that, strictly speaking, we don't *know* much of anything.

This doesn't mean that we are dumb, or that all our ideas about the world are useless, or that the only honorable course of action is to throw up our hands and throw in our lots with the Skeptics. "When one admits that nothing is certain," proposed the philosopher Bertrand Russell, "one must, I think, also add that some things are much more nearly certain than others." Those are sound words to live by. And yet, as the example of blind Hannah reminds us, we must also accept that we can't ascertain in advance which of the things we think we know will turn out not to be knowledge after all—will turn out, instead, to be wrong.

As that suggests, the idea of knowledge and the idea of error are fundamentally incompatible. When we claim to know something, we are essentially saying that we can't be wrong. If we want to contend with the possibility that we *could* be wrong, then the idea of knowledge won't serve us; we need to embrace the idea of belief instead. This might feel like an unwelcome move, since all of us prefer to think that we know things rather than "merely" believing them. That preference accords with the conventional view of knowledge and belief, in which the former is the loftier of the two concepts. In that view, knowledge is, you will recall, belief *plus* plus all the conditions the philosophers put on it, and all the faith that we ourselves put in it.

In the end, though, it is belief that is by far the broader, more complex, and more interesting category. It is, I will argue, the atomic unit of our intelligence—the thing that differentiates us from machines, and that lets us navigate the world as deftly as we do. But it is true (and not coincidental) that belief is also the atomic unit of error. Whether we wrongly think we can see or wrongly remember what we did on September 11, whether we are mistaken about pantyhose or mistaken about string theory, what we are ultimately wrong about is always a belief. If we want to understand how we err, we need to look to how we believe.

5.

Our Minds, Part Two: Belief

'Tis with our judgments as our watches, none
Go just alike, yet each believes his own.

—ALEXANDER POPE, "AN ESSAY ON CRITICISM"

On October 23, 2008, Alan Greenspan, the former chair of the Federal Reserve, appeared before a committee of the U.S. House of Representatives to testify about the financial crisis that had lately engulfed more or less the entire planet. Not surprisingly, the atmosphere was somber, and Henry Waxman, the California Democrat who chaired the committee, was not in the mood to pull his punches. "The Federal Reserve had the authority to stop the irresponsible lending practices" that had fueled the crisis, Waxman reminded those in attendance. But, he continued, "its long-time chairman, Alan Greenspan, rejected pleas that he intervene." Then he addressed Greenspan directly, and reproachfully: "those who trusted the market to regulate itself, yourself included, made a serious mistake."

This was hardly the kind of congressional welcome to which Greenspan was accustomed. Throughout his five terms as Fed chief, Greenspan had regularly been referred to as “the greatest central banker in history,” “the most powerful man in the world,” and, simply, “the maestro.” His fame extended beyond the United States (France awarded him the Legion of Honor; Great Britain made him an honorary knight), as well as beyond financial and political circles. In the words of the *Economist* magazine, Greenspan enjoyed “almost rock-star status” among ordinary Americans—a remarkable and somewhat baffling achievement for a famously tight-lipped financier in charge of a blindingly complex aspect of government. His autobiography, presciently titled *The Age of Turbulence*, reportedly sold for \$8.5 million dollars, second only to Bill Clinton’s memoir and tied with Pope John Paul II’s. When the book came out, in the summer of 2007—right around the time when the financial fault lines began to tremble—it topped both the *New York Times* and the Amazon.com bestseller lists.

By October 23, though, all that was in the past. The economy had been in bad shape for over a year, and in the spring of 2008, with the collapse of the global investment giant Bear Stearns, it had entered a virtual freefall. What began as a subprime mortgage crisis (triggered by the now-infamous practice of offering mortgages to people with limited or troubled credit histories) had broadened into a liquidity crisis, a credit crisis, a banking crisis, a currency crisis, a trade crisis—just about every kind of economic crisis you could name. In the United States, the stock market had fallen 37 percent since the start of the year. The American economy had lost 1.5 million jobs (a figure that would rise to over 5 million by early 2009), and the unemployment rate was marching toward double digits in the worst-hit states and sectors. Globally, the situation was even grimmer. The International Labor Organization predicted that, worldwide, between 18 and 50 million jobs would vanish into the maw of the crisis. Six months after Greenspan spoke, the Blackstone Group, a financial services company, reported that between 40 and 45 percent of global wealth had evaporated in under a year and a half.

For anyone with so much as a toe in the global economy, the crisis came as a massive financial and emotional shock—disturbing if you were lucky,

devastating if you weren’t. But for those who were charged with actually understanding and directing that economy, the collapse occasioned a massive ideological crisis as well. One of the primary jobs of economists is to create models of how financial systems work, and while those models are, by definition, simplifications and estimations as compared to the real thing, they are nonetheless supposed to be useful for making predictions, not to mention policy. (That’s why economists create them, after all.) As Waxman pointed out, Greenspan’s economic model was based on the premise that markets could be trusted to regulate themselves—and, as a corollary, that governments should not do so instead. Since Greenspan’s model had essentially been the global model for close to twenty years, the doctrine of market self-regulation had become all but holy writ. As Waxman put it, “trust in the wisdom of the markets was infinite.”

Then the markets imploded—and, with them, the model. As Greenspan told the committee, “the whole intellectual edifice collapsed.” He had, he continued, “found a flaw in the model that I perceived is the critical functioning structure that defines how the world works.” If that wasn’t clear enough, Waxman offered a blunt translation: “you found that your view of the world, your ideology, was not right.” It is a measure of how completely and publicly Greenspan’s model had failed that he was forced to concur. “Precisely,” he replied. “That’s precisely the reason I was shocked, because I had been going for forty years or more with very considerable evidence that it was working exceptionally well.” The demise of his doctrine that the market would always protect investors had left him, he said, “in a state of shocked disbelief.”

In that, Greenspan was hardly alone. One of the most striking features of the economic catastrophe was the sheer number of financiers wandering around in a state of stunned bewilderment, trying to grasp how their understanding of the world had served them so poorly. As the hedge fund manager Steve Eisman told the financial writer Michael Lewis (he of *Moneyball* and *Liar’s Poker* fame), being an investment banker in the early twentieth century was “like being a Scholastic, prior to Newton. Newton comes along, and one morning you wake up: ‘Holy shit, I’m wrong.’”

Still, one could argue—and many did—that Greenspan, at least, had

no business being quite so shocked. Over the years, countless people had challenged his deregulatory dogma, including (to name just a few) Joseph Stiglitz and Paul Krugman, both Nobel Prize-winning economists, and Brooksley Born, who was head of the Commodity Futures Trading Commission from 1996 to 1999. Born eventually became something of a Cassandra figure for the crisis, since she repeatedly called for regulating the market for derivatives, those ultracomplex financial products that eventually helped bring down the economy. Those calls were silenced when Greenspan, along with then-Treasury Secretary Robert Rubin and then-Secretaries and Exchange Commission Chair Arthur Levitt, took the extraordinary step of convincing Congress to pass legislation forbidding Born's agency from taking any action for the duration of her term. In a joint statement issued at the time, Greenspan defended the move on the grounds of "grave concern about this [proposed regulatory] action and its possible consequences." Merely discussing the option of government regulation, he asserted, could destabilize the markets and send capital surging out of the United States.

If Greenspan was reduced to "shocked disbelief" when the markets failed to regulate themselves and slid into chaos, it was not because he'd never been warned of the possibility. Nor was it because his own model had never been criticized (it had), or because alternative models had never been floated (they had). The problem, instead, was that his faith in the ability of markets to regulate themselves was, in Born's word, "absolutist." Greenspan was as figuratively invested in unregulated markets as the rest of us were literally invested in them. He had a model of how the world worked, and his confidence in it was all but immovable.

Actually, Greenspan had many thousands of models of how the world worked. He must have, because we all do. These models are our beliefs, and they cover everything from how we should invest our money to where we left our wallet. We believe in some of these models only tentatively—personally, I'm only about 50 percent sure of the whereabouts of my wallet right now—and some of them absolutely. But no matter how unshakably we believe in them, the models themselves can be shaken; that is what differentiates belief from the imaginary ideal of knowledge. Knowledge, as

we have seen, cannot make room for error, and therefore could not have failed Greenspan. But beliefs can, and his did. In fact, it failed us all.

This book is about our Greenspan moments: about what happens when our beliefs, including our most fundamental, convincing, and important ones, fail us. To understand how beliefs fail, though, we first need to understand how they work. And to understand that, we need to start with the most basic question of all: just what *is* a belief, anyway?

When we talk about beliefs in casual conversation, we usually mean our overt convictions about important matters: about religion or morality or propriety, politics or economics, ourselves or other people. These beliefs are explicit, in the sense that we are aware of having them, and can articulate and defend them if called upon to do so. Put differently, there is an experience associated with holding them. It felt like something for Alan Greenspan to believe in market self-regulation, just as it feels like something for you to believe in God, or in universal healthcare, or that your father-in-law dislikes you.

By contrast, when philosophers talk about belief (which they do often; it is an occupational hazard), they mean something markedly different. Or rather, they mean something *more*: they agree that our overt convictions about financial markets and so forth deserve to be called beliefs, but they think a lot of other things merit the term as well. Suppose you are reading this book in bed at midnight with the blinds drawn. Philosophers would say that your set of beliefs right now includes the following: that it is dark outside; that the sun will not rise for many more hours; that when it does, it will do so in the east; that the mattress underneath you is a solid object; that a flying saucer is not about to crash through your bedroom window; that you will wake up tomorrow as the same person you are today, and so on.*

* And on and on: one implication of the philosophical definition of belief is that, technically, the set of beliefs each of us holds is infinite. (If you believe there isn't a monster under your toddler's bed, you also believe there aren't two monsters under her bed, or three monsters, or . . . et cetera.) Needless to say, no one actually holds an infinite number of beliefs in the conscious mind. But the point of this definition of belief is that consciousness doesn't matter. What matters is that our everyday actions are grounded

What makes these additional beliefs seem so strange—what makes them seem, in fact, so unbelief-like—is that there is no experience associated with holding them. Believing that my mattress is solid doesn't feel like believing in God, largely because believing that my mattress is solid doesn't feel like *anything*. I am completely unaware of believing it. Left to my own devices, I'm extremely unlikely to characterize it as a belief, and, if I'm called upon to defend it, I will be both baffled as to why I should do so and at a loss as to how to go about it. In other words, almost everything we normally associate with the experience of believing—consciousness, conviction, emotion, explanation—is absent from these other, implicit beliefs.

Psychologically, then, the everyday concept of belief and the philosophical one differ from each other in the most salient way imaginable: in how we experience them. Functionally, however, they are virtually indistinguishable. Whether we are aware of our beliefs or not, they are all, like Greenspan's free-market philosophy, models of the world. In the literal sense, a model of the world is a map, and that's basically what beliefs are, too: mental representations of our physical, social, emotional, spiritual, and political landscapes. My explicit belief that my father-in-law dislikes me is crucial to my mental representation of my family, just as my implicit belief that my mattress is solid is crucial to my mental representation of my bedroom. Both serve the same maplike function of helping me figure out where I might or might not want to sit when I enter a certain room. Both, in other words, are necessary pixels in my picture of the world. Whether or not I can feel that pixel light up in my head is irrelevant. Think about what happens when I walk into an unfamiliar hotel room at night: in order to decide where to lie down, I need a mental image of the kind of things on which I can sleep, but I don't need to know that I have that image. The model of the world—the belief—is vital, but awareness of the belief is dispensable. In fact, lack of awareness is the norm. From the anticipated behavior of inanimate objects to the presumed identity of our parents to whether we can see a mountain chain or see at all, the vast

in an essentially limitless number of implicit convictions.

majority of our mental models are implicit: entirely unfelt, yet essential to how we make sense of ourselves and the world.

Just as implicit and explicit beliefs function in the same way, they also fail in the same way. However different they might seem to us under normal circumstances, in the moment of error, they are identical. Or rather, they *become* identical: the instant an implicit assumption is violated, it turns into an explicit one. Imagine for a moment a scene worthy of a Marx Brothers movie. It's nighttime, I emerge from the bathroom in my pajamas, pick up my book, lie down on my bed—and, wham, fall straight through the mattress onto the floor. Should this exceedingly unlikely scenario somehow transpire, three things will have collapsed. The first is my mattress. The second is my belief in the solidity of that mattress. The third—and here is the point I am trying to make—is the implicitness of that belief. If I find myself sprawled on the floor, all of my previously unconscious convictions about mattresses will suddenly surge into consciousness. In the moment of error, our implicit beliefs are simultaneously contravened and revealed.

Once we acknowledge that implicit assumptions and explicit convictions are just subsets of the single category of belief, we can go further and acknowledge that even the distinction between them is suspect. All of us have plenty of beliefs that we are somewhat aware of holding, or that we can become aware of holding when necessary. (I don't believe anything about the location of my sunglasses in December, but come June, when I can't find them, I believe that my sister borrowed them.) Still, however tenuous this distinction might be, it is relevant to this book in at least one respect. While many of our beliefs fall somewhere in the middle of the implicit-explicit spectrum, it is those that lie at the extreme ends that collapse most spectacularly in the event of error. If anything can rival for sheer drama the demise of a belief that we have adamantly espoused, it is the demise of a belief so fundamental to our lives that we never even registered its existence.

Our beliefs, then, are models of the world—but they aren't just models for models' sake. Like economic models, our mental ones exist to help us make predictions and policies. In the words of William James, beliefs "are re-

ally rules for action." One obvious corollary to this is that our beliefs have consequences. I don't mean that being *wrong* about them has consequences, although that's also true: as we just saw, a flawed belief helped eradicate nearly half the world's wealth. I mean that simply *holding* a belief can have consequences.

Consider another financial example. Since 1992, CalTech and MIT have been collaborating on a project known as the Laser Interferometer Gravitational-Wave Observatory. The observatory is the single most expensive project ever funded by the National Science Foundation. It took ten years and \$300 million to build, and costs \$30 million per year to operate. Its mission is to find gravitational waves ("ripples in the fabric of space-time [that] are produced by violent events throughout the universe," says the NSF, if that helps), and it had better succeed, because the existence of those waves is posited by nothing less than the general theory of relativity. As the science writer Margaret Wertheim has noted, this is an instance where "belief has brought into being a half-billion-dollar machine."

A half-billion dollars is a bundle of money, but pricey machines are just the tip of the iceberg. Housing bubbles, holy wars, the 1964 Civil Rights Act, the lunar landing: all of these came about as the result of belief. Plainly, then, our beliefs don't just have financial and material consequences. They also have political, legal, intellectual, cultural, and even corporeal ones. And, crucially, they have emotional and psychological consequences as well. Again, I'm not talking about the emotional consequences of being wrong about a belief, a topic we'll get into later. I'm talking about the emotional consequences of merely believing it—the way that building a gravitational-wave observatory is a consequence of believing in general relativity (whether or not it is correct), and investing in the stock market is a consequence of believing that it is sound (whether or not it is).

Some of the emotional consequences of our beliefs are pretty straightforward. If you believe that your true love is nervous during dinner because he plans to propose to you over dessert, you will be excited and happy; if you believe that he is nervous because he plans to break up with you, you will be anxious and upset. But examples like that don't get at either the scope or the significance of the psychological repercussions of our beliefs. To grasp that

scope, you have to grasp the scope of belief itself. Our models of the world extend beyond markets and mattresses and the general theory of relativity, into a kind of General Theory of Us: whether we think we are attractive and intelligent, competent or inept, better or worse off than other people; whether we think our parents loved us; whether we think a God is watching over us; whether we think we are basically safe and cared for in the world.

Convictions like these organize our idea of who we are, as well as how we relate to our environment. As that suggests, and as we'll see throughout this book, our beliefs are inextricable from our identities. That's one reason why being wrong can so easily wound our sense of self. And it's also why psychotherapy often focuses on helping people to examine—and, when necessary, change—their beliefs about themselves and others. Regardless of whether those beliefs are conscious or unconscious, regardless of whether they are right or wrong, they determine how we feel and how we behave every day of our lives.

Here is what we have learned so far in our quest to understand beliefs: they are models of the world; they help us take action; and accordingly, they incur consequences. Good enough—except why, then, do we have so many beliefs that we will never be able to act on? This is sometimes known as the problem of distal beliefs. "Distal," in this case, means far from the self; a distal belief is one that pertains to things remote from us in time or space or relevance. If we believe in the soundness of markets, the solidity of mattresses, or the existence of God, those beliefs will guide our actions in the world. But what about the belief that string theory is right, that South Africa's AIDS policy is wrong, and that Alpha Centauri C is the closest star to the earth? As we'll see in Chapter Seven, merely espousing such beliefs might be socially prudent or advantageous, regardless of whether they are relevant, or even right. But unless you are a physicist or a public health expert or an interstellar traveler, they will not enable you to take any action in the world.

Why, then, do we bother having distal beliefs? One way to answer this question—certainly the most fun way—is to think about sex. The reason

we have a sex drive is to ensure that we reproduce, but the vast majority of our sexual activity does not result in offspring. We simply have an instinct to copulate, with the consequence that we have enough babies and (evolutionarily speaking) way more than enough sex. For at least a century, psychologists and philosophers have suggested that our urge to explain the world is analogous to our urge to populate it. Like making babies, they argue, making theories is so crucial to our survival that we have a natural drive to do so—what William James called a “theoretic instinct.” This is the impulse I gestured toward in the last chapter: the one that compels us to generate stories and explanations all the time, even at the risk of being featured in a magazine called *Modern Jackass*.

It’s easy to see why a theory drive would be evolutionarily advantageous. Imagine that you are your own earliest ancestor, trying to make your way in the world some 200,000 years ago. Somehow, you have to figure out that shaking a certain kind of tree will make edible fruit fall to the ground. You have to learn that berries of specific shapes and colors are nourishing, while other very similar berries can kill you. Upon hearing a rustling in the bushes, you have to be able to infer—pretty damn quickly—the presence of a predator, or of dinner. In other words, you must be extraordinarily adept at guessing what’s going on in your environment and why. This is precisely the skill set covered by theorizing, and its utility has not diminished over time. While my concerns about predators might be greatly diminished these days, I now need to be able to determine whether the stranger striding toward me wants to ask me for a light or relieve me of my wallet, whether the explosions

* At the risk of further muddying some already murky lexical waters, I use the words “belief” and “theory” almost interchangeably in this book. Some thinkers have called for a distinction between the two terms, arguing that theories are more explicit, more developed, or more explanatory than beliefs. But such distinctions are largely untenable, or at any rate extremely slippery. For instance, try applying them to the statement “I believe in God.” If you are a five-year-old, that statement almost certainly represents a belief; a not very developed, probably not very explanatory, and possibly not even very explicit conviction. On the other hand, if you are the pope, it is clearly a theory. Moreover—and more important for my purposes—none of the differences we can posit between beliefs and theories have any bearing on their function. Theories, like beliefs, exist to represent the world around us.

outside are a threat to my life or merely a fireworks display, and whether or not my two-year-old just ate the Lego that is suddenly nowhere to be seen. Theorizing, like fornicating, is of timeless use to our species.

The evolutionary utility of theorizing helps explain why we have such a hyper-abundance of beliefs—including those, like distal beliefs, that serve no obvious purpose. Just as our reproductive instinct somehow produced Paris Hilton videos and our language instinct somehow produced Proust, our theorizing instinct has long since exceeded the barebones requirements of survival. Because we needed to be able to theorize about some things (unfamiliar berries, rustling bushes), we wound up able to theorize about everything. And we do. How supernovas are formed, why autism is on the rise, what was going on with that married couple at the cocktail party, why we like one pair of pantyhose more than another: there is virtually no subject—no matter what domain of life it concerns, how pressing or trivial it may be, and how much or little we know about it—that is not suitable fodder for our theory-happy minds.

The evolutionary urgency of theorizing also helps explain why we form beliefs both constantly and unconsciously. Of course, we are capable of theorizing intentionally, too, and we do so all the time—both informally (as when we spend happy hour trying to figure out why our boss was in such a foul temper that afternoon) and formally (as when we spend our career trying to figure out what causes cancer). What we aren’t capable of doing is *not* theorizing. Like breathing, we can ignore the belief-formation process or control it—or even refine it—but whatever we do, it will keep on going for as long as we keep on living. And with good reason: if we want to eat dinner rather than be dinner, we are well served by a process so rapid and automatic that we don’t need to waste time deliberately engaging it.

As with our perceptual processes, this automatic theorizing generally careens into consciousness only when something goes wrong. For instance: not long ago, I arranged to meet an interview subject for coffee in Manhattan. When I walked into the café and she stood up to introduce herself, I had the common yet always somewhat startling experience of realizing that she looked nothing like what I had expected. What was strange about this experience—what is always strange about such experiences—is that, prior

to meeting this woman, I had no idea that I had any mental image of her at all. And yet, somewhere in the back of my mind, I had generated a picture of her, without any conscious awareness of doing so. Moreover, this process must have been quite sophisticated, since when I went back and thought about it, I realized that not only could I describe the person I had expected to see, I could pinpoint some of the factors that led me to draw my mistaken portrait: a name I associated with a certain era and ethnicity, a scrap of information that suggested a certain aesthetic, and so forth. In short, some very sophisticated theory making was going on in my mind, entirely without my awareness. This is true for all of us, all the time. Below the level of conscious thought, we are always amassing information from our environment and using it to add to or rearrange our model of the world.

As with the much-debated language instinct, the theorizing instinct is, itself, just a theory. No one knows if our capacity to generate hypotheses about the world is truly hard-wired. We do know, though, that it kicks in very early. For instance, there is suggestive evidence that babies as young as seven months are already theorizing about basic physical properties such as gravity. That might seem hard to believe, but tack on a few more months and you've got a toddler—and toddlers are infamous theorists. Armed with not much more than an insatiable drive for physical exploration and an enthusiasm for the word "why," the average two-year-old seems determined to take on the whole world. And that's exactly what he or she is doing. It is in childhood, after all, that our environment is most mysterious to us, and most urgently in need of mapping. The Berkeley psychologist Alison Gopnik has even posited that the theory drive exists specifically for early childhood, although it operates throughout our lives—just as the sex drive exists specifically for our fertile years, even though we are sexual long before and long after.

Whether the idea of a theory instinct turns out to be sound biology or just apt analogy, belief formation is clearly central to our species. Its survival value aside (as if survival were the kind of thing we could set aside), it defines the way we inhabit our environment, inarguably for the better. I invoke again the analogy to sex and language: it is good to make babies and shout warnings, but it is *really* good to make love and read Shakespeare.

So too with theorizing: the instinct is about staying alive, but the excess is about living. It enriches our everyday life (those happy-hour debriefs about the boss) and enables our most extraordinary achievements (that hoped-for cure for cancer). Without it, we would be bereft of virtually all our hallmark human endeavors: religion, science, and storytelling; curiosity, exploration, and discovery. With it, we are able to venture beyond the known world and toward the terrain of all that is hidden from us—the past and the future and the far-flung, the cloaked machinery of nature, other people's inner lives. It is the gift that beckons us beyond the eked-out existence of mere survival.

There is, predictably, a problem. Although we are highly adept at making models of the world, we are distinctly less adept at realizing that we have made them. As I suggested in my discussion of perception, our beliefs often seem to us not so much constructed as reflected, as if our minds were simply mirrors in which the truth of the world passively appeared.

Psychologists refer to this conviction as naïve realism. Naïve realism is an automatic tendency, not an intentional philosophical position. If you actually believed that the world was precisely as you experienced it, you would be an intentional naïve realist, but as far as I know, no one has seriously subscribed to this position in the entire history of humanity. Even the most impassioned realist, the kind that regards relativists as dangerous lions from Planet France, recognizes that our experience of the world is not identical with the world itself. To take only the simplest examples: there is no such thing as red if you are a bat, and no such thing as loud if you are a rock, and (so far as we currently know) no such thing as triumph or regret if you are a Shetland pony. Color and sound and emotion are all central to how we experience and make sense of the world, but they can't inhere in the world itself, because they cease to exist the moment you take minds out of the picture. Conversely, plenty of things exist in the world that the human mind cannot experience directly: the infrared spectrum, the structure of molecules, and very possibly a dozen or so extra dimensions.

With good reason, then, there are no proponents of naïve realism. But that doesn't mean there are no naïve realists. On the contrary, there are

tons of them—starting, research suggests, with everyone under the age of four. Very young children seem to believe, truly and ardently, that our minds and the world never diverge from each other. As a result—and this is why naïve realism matters so much to this book—they think we can't believe things that are wrong.

We know this about little kids courtesy of one of the classic experiments of developmental psychology. The experiment is known as the false belief test, or, informally, as the Sally-Ann task, after the most famous of its many variations. If you count among your acquaintances a child under the age of four, you can replicate it yourself—and you should, because until you've seen the results firsthand, it's hard to grasp their weirdness, and harder still to believe them. All you need to do is stage a simple puppet show. One character (Sally, in the classic version) takes a candy bar, puts it in a basket, closes the lid, and leaves the room. As soon as she disappears, another character (Ann) takes the candy bar out of the basket and hides it in a cupboard. Now ask the child this: when Sally returns, where will she look for her candy bar?

If the child you've enlisted for this experiment is your own, you already know that kids of this age are stupendous thinkers. They speak and understand their native language (or languages) with ease, and add words and concepts with a rapidity that is the envy of every adult. They are insatiably curious, highly attentive to the world around them, and capable of impressive feats of memory and concentration. They understand cause and effect and can reason about when, why, and how things happen. They make inferences about the world around them that, even when wrong, display a remarkable attentiveness to their environment. They play games, not to mention invent them. They negotiate complex social interactions, and they understand that different people can have different needs, desires, and emotions. Depending on the family they were born into, they may already be reading about star-bellied sneetches or learning to play hockey or studying the violin. And yet, without fail, these same brilliant children will confidently report that Sally will look for her candy bar not in the basket where she put it, but where it actually is: in the cupboard where it was hidden while she was out of the room.

To adults (or, for that matter, six-year-olds), this answer is baffling. We understand that there is no way Sally could know the real whereabouts of her candy bar—literally no mechanism by which she could acquire that knowledge—because she wasn't around to witness Ann relocating it. But this experiment suggests that young children don't care about such mechanisms. For them, it seems, the world and the mind enjoy an automatic correspondence: Sally thinks the candy bar is in the cupboard because the candy bar *is* in the cupboard. Adults, whether they are realists or relativists, recognize that the mind contains a kind of personalized representation of reality—the world as rendered by you or me or Sally. Kids, by contrast, seem to think the mind contains a *replica* of reality: the world as rendered by Xerox. Apparently, they don't yet understand that beliefs about the world can be at variance with the world itself.

If the Sally-Ann test doesn't seem like conclusive proof of this claim, consider the following, even more astonishing variant of it. In this version, children are shown a box with a picture of candy on the front and asked what they think is inside. Reasonably enough, they say "candy." When they open it, however, they find (presumably to their disappointment) that it is full of pencils. Here's the astonishing part: if you then ask the children what they thought was in the box before they opened it—that is, about twenty seconds earlier—they will insist that they thought it contained pencils. What children maintain about the imaginary Sally they also maintain about themselves: that their beliefs about the world cannot deviate from the world as it really is.

This faith in the perfect accuracy of our beliefs is fleeting. By the age of five, virtually all children can pass the Sally-Ann test with ease. In coming

* Remember the Pessimistic Meta-Induction from the History of Science? Here it is in action. While I was writing this book, the once uncontroversial claim that young children can't grasp the existence of a gap between the mind and the world suddenly came up for debate. Recent evidence from infancy experiments (where eye gaze is used as a measure of babies' beliefs) suggests that children might understand more about false beliefs than psychologists previously thought. Although three- and four-year-olds still reliably fail the Sally-Ann task, it now appears that fourteen-month-olds can pass it. It remains to be seen how these findings will be reconciled.

to do so, these children have acquired what developmental psychologists call "representational theory of mind." That is, they've figured out what a mind is, at least in general terms—not a photocopy of reality but a private and somewhat idiosyncratic means for making sense of the world—and they've figured out that everybody has one. This changed understanding leads to striking new insights: that beliefs about the world can be at odds with the world itself; that my beliefs can be at odds with yours; that other people don't necessarily know everything I know; and, conversely, that I don't necessarily know everything other people know.

These insights seem so obvious to adults that it is easy to overlook their importance. The ability to grasp what minds do, to understand that people can hold beliefs that are mistaken or different from our own, subtends a vast swath of mature thinking. It allows us to "read minds," not as psychics use the term but as psychologists do—to infer people's thoughts and feelings based on their words, actions, or circumstances. Rebecca Saxe, a neuroscientist at MIT and one of the leading contributors to our understanding of the brain structures underlying theory of mind, offers the example of *Romeo and Juliet*. As audience members, we know that the seemingly lifeless Juliet is not actually dead, as Romeo believes, but has merely taken a sleeping potion. But if we didn't have theory of mind, we wouldn't be able to set aside our own knowledge and see the scene as Romeo does—and so we wouldn't understand why he kills himself. The false belief on which the whole tragedy turns would be completely lost on us.

So, too, would entire expanses of the social landscape. Without theory of mind, we wouldn't be able to register the subtleties of a flirtation, recognize our accidental offenses against a friend, or foresee that coming home two hours late might alarm and anger our family. As these examples suggest, theory of mind is vital to our emotional, intellectual, and moral development. (Tragically, we have some idea of how compromised we would be without it, because its absence or diminution is characteristic of people with autism and Asperger's syndrome.)

* Another variant of the Sally-Ann task provides a particularly poignant illustration of the difference between healthy and autistic children's ability to understand other

Once you acquire theory of mind, there is no going back; barring serious brain injury, you will never fail the Sally-Ann test again. But the attraction of naïve realism never wholly fades. Granted, we come to understand, in the abstract, that our beliefs can be skewed by any number of factors, ranging from the silent nudgings of self-interest to the limits of omniscience—the fact that, like Sally, sometimes we just aren't in the right room at the right time. When it comes to our specific convictions about the world, however, we all too easily lapse back into the condition of toddlers, serenely convinced that our own beliefs are simply, necessarily true.[†]

Why do we do this? The most obvious answer is that we're so emotionally invested in our beliefs that we are unable or unwilling to recognize

people's minds. In this version, the experimenter shows the child a Polaroid camera, explains what it does, takes some sample pictures, and allows the child to play with it until he or she is familiar with how it works. Then the puppet show proceeds as in the original experiment—except that, when Sally puts the candy bar in the basket, the experimenter takes a picture of it there. At the end of the show, after Sally has left the room and Ann has moved the candy bar to the cupboard, the child is asked not where Sally thinks the candy bar is, but where it will appear in the photograph. Although this test and the original Sally-Ann experiment seem identical in the determination that children are asked to make, neuroscientists have shown that we use different parts of the brain to reason about minds than about objects. Healthy children find the Polaroid version of the false-belief task *harder* than the original one: they continue to fail it for some months after they've begun to reason correctly about other people's beliefs. Autistic children, by contrast, can pass the false-belief test when it involves Polaroid pictures, but not when it involves other people's minds. The workings of the former, a mechanical object, are transparent to them. It is the latter that is the black box.

† Our habit of treating many of our beliefs like facts can be seen, in somewhat sidelong fashion, in the specialized way we use the expression "I believe." When I flag a statement as a belief, I'm not emphasizing the depths of my conviction; I'm taking pains to convey my doubt. Imagine that you and I are at a party when a distant acquaintance walks through the door. You nudge me and ask me to remind you of his name. If I say, "His name is Victor," there's a good chance that you'll exclaim, "Victor! Good to see you again." But if I say, "I believe his name is Victor," you will probably have the sense to hold back. You will rightly hear, in that "I believe," an implicit caveat: "this is *just* a belief; I could be wrong." We attach no such caveats when we feel no uncertainty, which is why we say "Rudy Giuliani is a reactionary control freak," not "I believe that Rudy Giuliani is a reactionary control freak."

them as anything but the inviolable truth. (The very word "believe" comes from an Old English verb meaning "to hold dear," which suggests, correctly, that we have a habit of falling in love with our beliefs once we've formed them.) There's a lot to be said for that answer, and much of it *will* get said in the coming pages. For now, though, I want to propose another, less obvious theory about why we act like our beliefs are necessarily true—which is that we are logically obliged to do so. Philosophers have a name for this theory, but, unfortunately, it's a name that only a philosopher could love: the First Person Constraint on Doxastic Explanation. "Doxastic" means "pertaining to beliefs"; that strange syllable, *dox*, is the same one that shows up in words like "orthodox" ("believing correctly") and "paradox" ("contrary beliefs"). In lay terminology, the phrase means that each of us has limited options for how to explain why we believe the things we do.

I'm going to jettison this cumbersome name and (for reasons that will become obvious in a moment) call this idea the 'Cuz It's True Constraint. Here's how it works. Let's say I believe that drinking green tea is good for my health. Let's also say that I've been drinking three cups of green tea a day for twenty years, that I come from a long line of green tea drinkers, and that I'm the CEO of a family-owned corporation, Green Tea International. An impartial observer would instantly recognize that I have three very compelling reasons to believe in the salutary effects of green tea, none of which have anything to do with whether those effects are real. First, having ingested vast quantities of it, at least partially in the conviction that I was boosting my chances at a long and healthy life, I'm going to be resistant to any suggestion that all that tea had zero effect on me (or, worse, a deleterious one). Second, because my allegiance to green tea is part of an entrenched and presumably sacrosanct family tradition, questioning it could seriously damage my most intimate relationships, not to mention my share of the family fortune. Finally, I have staked my financial and professional status on the belief that green tea is good for one's health.

In short, I have powerful social, psychological, and practical reasons to believe in the merits of green tea. The gist of the 'Cuz It's True Constraint is that I myself can't believe that these reasons contribute in any significant way to my conviction that green tea is good for me. Instead, I must believe

that this conviction is based on the facts: in this case, on the physical (rather than emotional, financial, or familial) benefits of green tea. In other words, I must believe that I believe it *it's true*. As the philosopher Ward Jones said, "It simply does not make sense to see myself as both believing that P is true"—where "P" stands for any proposition—"and being convinced that I do so for reasons having nothing to do with P's being true."

Viewed in a certain light, the 'Cuz It's True Constraint can appear self-evident, or even circular. *Of course* we have to think our beliefs are true: that's what it means to believe them. Fair enough. But one of the strengths of philosophy lies in looking closely at the self-evident—and when you look closely at the 'Cuz It's True Constraint, you see the origins of some of the most important aspects of our relationship to wrongness. Specifically, you begin to see why we are so convinced that our own beliefs must be right, and why we feel no need to extend that assumption to other people.

So let's look closely. The 'Cuz It's True Constraint has several stipulations, the first of which is that it applies only to beliefs I currently hold. I can readily concede that beliefs I *used* to hold weren't based on the facts—that, say, my conviction about adulterers burning in eternal hellfire was just a product of my evangelical upbringing, or that my stint with the International Socialist Organization was just a way to rebel against my conservative parents. What's more, once I've rejected a belief, I can often *only* perceive the self-serving reasons I believed it, and can no longer recognize any evidence for it as rationally compelling.

The second stipulation of the 'Cuz It's True Constraint is that it applies only to specific beliefs—not to my entire set of beliefs, nor to how I feel about the nature of believing in general. As I already mentioned, most of us can acknowledge, in the abstract, that beliefs are influenced by all kinds of non-objective factors. We can even go further and admit that, at this very moment, some of our own beliefs are doubtless swayed by such factors as well. It is only when we are confronted about a specific and active belief that the constraint kicks in. Ask me whether I think my beliefs in general are affected by personal biases and I'll say sure. Ask me about the fidelity of my girlfriend or the safety of my health regimen or the accuracy of the data I just published in that prominent journal, and—ah, well, I assure you that I

believe in *those* things not because they are comforting or convenient, but because, by God, they are true.

Finally, consider the most important stipulation of the 'Cuz It's True Constraint. This is the one suggested by the "first person" part of its proper name: it applies only to our own beliefs, not to those of other people. Nothing about the constraint prevents me from thinking that Ellen believes in God to ameliorate her fear of death, or that Rudolf opposes gun control because his father sits on the board of the NRA, or that you believe behaviorism is baloney because your entire tenure committee does, too. On the contrary, we impute biased and self-serving motives to other people's beliefs all the time. And, significantly, we almost always do so pejoratively. If I suggest that the CEO of Green Tea Incorporated stands to gain financially by believing in the benefits of green tea, I'm at the very least implying that she isn't qualified to judge the truth of her belief—and, more likely, I am implying that there *is* no truth to her belief. In other words, if we want to discredit a belief, we will argue that it is advantageous, whereas if we want to champion it, we will argue that it is true. That's why we downplay or dismiss the self-serving aspects of our own convictions, even as we are quick to detect them in other people's beliefs.

Psychologists refer to this asymmetry as "the bias blind spot." The bias blind spot can be partly explained by the Lake Wobegon Effect, that endlessly entertaining statistical debacle whereby we all think that we are above average in every respect—including, amusingly, impartiality. But a second factor is that we can look into our own minds, yet not into anyone else's. This produces a methodological asymmetry: we draw conclusions about other people's biases based on external appearances—on whether their beliefs *seem* to serve their interests—whereas we draw conclusions about our own biases based on introspection. Since, as we've seen, much of belief-formation neither takes place in nor leaves traces in conscious thought, our conclusions about our own biases are almost always exculpatory. At most,

* This is a logical fallacy. The fact that I benefit from a given belief might raise questions about my capacity to assess it objectively, but it doesn't bear on the truth of the belief itself. After all, people benefit from true beliefs as well as false ones.

we might acknowledge the existence of factors that could have prejudiced us, while determining that, in the end, they did not. Unsurprisingly, this method of assessing bias is singularly unconvincing to anyone but ourselves. As the Princeton psychologist Emily Pronin and her colleagues observed in a study of the bias blind spot, "we are not particularly comforted when others assure us that they have looked into their own hearts and minds and concluded that they have been fair and objective."

So we look into our hearts and see objectivity; we look into our minds and see rationality; we look at our beliefs and see reality. This is the essence of the 'Cuz It's True Constraint: every one of us confuses our models of the world with the world itself—not occasionally or accidentally but *necessarily*. This is a powerful phenomenon, and it sets in motion a cascade of corollaries that determines how we deal with challenges to our belief systems—not, alas, for the better.

The first such corollary is the Ignorance Assumption. Since we think our own beliefs are based on the facts, we conclude that people who disagree with us just haven't been exposed to the right information, and that such exposure would inevitably bring them over to our team. This assumption is extraordinarily widespread. To cite only the most obvious examples, all religious evangelism and a good deal of political activism (especially grassroots activism) is premised on the conviction that you can change people's beliefs by educating them on the issues.

The Ignorance Assumption isn't always wrong; sometimes our ideological adversaries *don't* know the facts. But it isn't always right, either. For starters, ignorance isn't necessarily a vacuum waiting to be filled; just as often, it is a wall, actively maintained. More to the point, though, the Ignorance Assumption can be wrong because *we* can be wrong: the facts might contradict our own beliefs, not those of our adversaries. Alternatively, the facts might be sufficiently ambiguous to support multiple interpretations. That we generally ignore this possibility speaks to the powerful asymmetry of the Ignorance Assumption. When other people reject our beliefs, we think they lack good information. When we reject their beliefs, we think we possess good judgment.

When the Ignorance Assumption fails us—when people stubbornly

persist in disagreeing with us even after we've tried to enlighten them—we move on to the Idiocy Assumption. Here, we concede that our opponents know the facts, but deny that they have the brains to comprehend them. This assumption can be a narrow judgment, applied to a specific person on a specific issue, or it can be a sweeping assessment of any individual or group we regard as the opposition. In the course of working on this book, I spoke with a left-wing lawyer who described growing up in a politically progressive environment and then attending a liberal (in every sense) arts college. As a consequence, she told me, "it wasn't until I went to Yale Law School that I met people I disagreed with, ideologically, who were incredibly smart. This is going to sound ridiculous, but it wasn't until then that I realized that conservatives could be intelligent." It would be nice if that statement *did* sound ridiculous, but we hear variations on it so often that this one scarcely registers as surprising (except, perhaps, in its forthrightness). Think, for instance, of the countless times we say things like "what kind of idiot could actually believe . . ."

One of the most common answers to that question is: the wicked kind. This is the Evil Assumption—the idea that people who disagree with us are not ignorant of the truth, and not unable to comprehend it, but have willfully turned their backs on it. The Evil Assumption has a longstanding relationship with religion, where "unbeliever" is all but synonymous with "evildoer." But it is almost equally common in politics. In *The Prelude*, the poet William Wordsworth acidly described the French Revolution as a cause that, ostensibly, "no one could stand up against, / who was not lost, abandoned, selfish, proud, / mean, miserable, willfully depraved, / hater per-verse of equity and truth." (Wordsworth was largely condemning his own doctrinaire past; the poem is subtitled *Growth of a Poet's Mind*.) As those lines suggest, clashes of belief that engender violent conflict are especially good at provoking the Evil Assumption. And, conversely, the Evil Assumption is especially good at provoking violent conflict. (If beliefs have consequences, consider the likely consequences of believing that those who disagree with you are wicked.) But those who disagree with you don't have to reject your God or threaten your life for you to conclude that they are depraved. These

days, you can't punch the scan button on your radio without coming across an invocation of the Evil Assumption: hosts or guests or callers describing their ideological opponents as morally depraved reprobates bent on the destruction of civilization as we know it.

That's quite a charge to lay at the feet of people who don't agree with us. And yet it has a certain dark logic, given our tendency to confuse our models of reality with reality itself. Think about the accusation that people who disagree with us "don't live in the real world." What we really mean is that they don't live in our model of the world; they don't share our vision of how things are. By failing to see the world as we do, they actually *are* undermining its reality and threatening its destruction—at least, unto us. But, of course, we are doing the same to them. Implicitly or explicitly, we are denying that they possess the same intellectual and moral faculties that we do—and denying, too, the significance and value of their life experiences, from which, inevitably, many of their beliefs arise.

Of these three assumptions—the Ignorance Assumption, the Idiocy Assumption, and the Evil Assumption—the last is the most disturbing. But the first is the most decisive. We assume that other people are ignorant because we assume that we are not; we think we know the facts, and (as the 'Cuz It's True Constraint mandates) we think those facts determine our beliefs. Put differently, we think the evidence is on our side. It is almost impossible to overstate the centrality of that conviction to everything this book is about, which is why we are going to turn next to the topic of evidence. Our faith in our own reading of the facts undergirds our certainty that we are right, our shock when we turn out to be wrong, and our willingness to deny the perspicacity, intelligence, and moral worth of everyone who disagrees with us.

What is alarming is how naturally this cascade of assumptions comes to us—and not just to the extremists in our midst. I wouldn't characterize Alan Greenspan as a rabid ideologue, and I'd venture to guess that he doesn't routinely use the word "evil" to describe people who disagree with him about economic policy. But he *did* regard such people as dangerous, and (as countless real extremists have likewise done) he did seek to silence them.

Those are not the actions of a man who has contemplated the possibility that his own model could be in error—or that, this time, his beliefs could be the dangerous ones.

This is the grim flip side of our passion for inventing theories. Like toddlers and tyrants, we are quick to take our own stories for the infallible truth, and to dismiss as wrongheaded or wicked anyone who disagrees. These tendencies are most troubling for the way they fuel animosity and conflict. But they are also troubling because they make it extremely difficult to accept our own fallibility. If we assume that people who are wrong are ignorant, or idiotic, or evil—well, small wonder that we prefer not to confront the possibility of error in ourselves.

6.

Our Minds, Part Three: Evidence

ROSENCRANTZ: That must be east, then. I think we can assume that.

GUILDENSTERN: I'm assuming nothing.

ROSENCRANTZ: No, it's all right. That's the sun. East.

GUILDENSTERN (*looks up*): Where?

ROSENCRANTZ: I watched it come up.

GUILDENSTERN: No . . . it was light all the time, you see, and you opened your eyes very, very slowly. If you'd been facing back there, you'd be swearing *that* was east.

—TOM STOPPARD, *ROSENCRANTZ AND GUILDENSTERN ARE DEAD*

In 1692, Judge William Stoughton, one of the most venerable lawmakers in Colonial America, found himself facing an unusual procedural question: Should visitations by evil spirits be admitted as evidence in a court of law? Stoughton was presiding over the Salem witch trials and—unfortunately for the 150 people who were imprisoned and the nineteen who were hanged by