BLINDSPOT
The Hidden Biases of Good People

Mahzarin R. Banaji
and Anthony G. Greenwald

DELCORTE PRESS / NEW YORK
The sailor cannot see the North
—but knows the Needle can—

EMILY DICKINSON, in a letter to a mentor, T. W. Higginson, seeking an honest evaluation of her talent (1862)
Cover

eBook Information

Title Page

Copyright Page

Dedication

Epigraph

Preface

1 Mindbugs

2 Shades of Truth

3 Into the Blindspot

4 “Not That There’s Anything Wrong with That!”

5 Homo Categoricus

6 The Stealth of Stereotypes

7 Us and Them

8 Outsmarting the Machine

Appendix 1 Are Americans Racist?

Appendix 2 Race, Disadvantage, and Discrimination

Acknowledgments

Notes

References

About the Authors
LIKE ALL VERTEBRATES, YOU HAVE A BLIND SPOT IN EACH EYE. THIS is a small region where the optic nerve attaches to the retina. Because there are no light-sensitive cells in this region, light reaching your blind spot (scotoma) has no path to visual areas of your brain.

Nevertheless, you can “see” your own blind spot by looking at the plus sign in the middle of the rectangle below with just one eye open. (You may need to take your glasses off.) Starting with the page about a foot in front of your nose, bring it closer while focusing on the plus sign. One of the black discs (the one on the same side as your open eye) will disappear at some point when the page is about six inches away, and will reappear as you bring the page closer still. The moment of disappearance tells you when light from that disc is falling on your open eye’s blind spot. Here’s a bonus: If you shift your open eye to look at the still-visible disc on the other side, the plus sign will disappear!

You may have also noticed a strange occurrence at the spot where the disc lies. When the disc disappeared, it left no blank spot in the grid background. Your brain did something interesting—it filled in the scotoma with something that made sense, a continuation of the same grid that you could see everywhere else in the framing rectangle.

A much more dramatic form of blindness occurs in the pathological condition called blindsight, which involves damage to the brain’s visual cortex. Patients with this damage show the striking behavior of accurately reaching for and grasping an object placed in front of them, while having no conscious visual experience of the object. If you placed a hammer before the patient and asked, “Do you see something before you?” the patient would answer, “No, I don’t,” but when asked to just reach for and grasp the hammer, she would do so as accurately as if she could see. This happens because the condition of blindsight leaves intact subcortical retina-to-brain pathways that suffice to guide behavior, even in the absence of the conscious experience of seeing the hammer.

We understand both the retinal blind spot and clinical blindsight as metaphors that help in understanding thoughts and feelings of which we are not aware, but which nevertheless guide our behavior. The hidden-bias blindspot of this book’s title has some features of both the retinal blind spot and blindsight. Like the retinal blind spot you experienced, we normally have no awareness of hidden biases. And like the dramatic phenomenon of clinical blindsight, hidden biases are capable of guiding our behavior without our being at
all aware of this guidance.

Hidden biases of the sort we find interesting can cause us to judge and act toward others in favorable or unfavorable ways that come from unrecognized feelings and beliefs about the groups to which people belong. In talking with many people about hidden biases, we have discovered that most find it simply unbelievable that their behavior might be guided in this fashion, without their awareness. Our main aim in this book is to make clear why scientists now regard this blindspot as fully believable. Convincing readers of this is no simple challenge. How can we show the existence of something in our own minds that is hidden from our own awareness?

Just as the rectangle with the two black discs allows us to “see” the otherwise hidden retinal blind spot, a device called the Implicit Association Test has enabled us to discover the contents of hidden-bias blindspots. Even better than the demonstration of the retinal blind spot, which allows us to know that we have a blind spot but not much more, the Implicit Association Test (IAT for short) lets us look into the hidden-bias blindspot and discover what it contains.

It is not often that scientists can learn something about their own minds while doing their research, but we have been the beneficiaries of such experiences. In addition to being a scientifically valuable device that we and many others have used in research for the past seventeen years, the IAT has also confronted us personally with some unwelcome, insight-provoking experiences—bringing us face-to-face with the contents of our own blindspots. In this book we give readers the opportunity to likewise experience the contents of their own blindspots.

Here is a brief overview of what’s to be found in this book.

We begin with “Mindbugs” (Chapter 1), which sets the stage for understanding hidden-bias blindspots. In it, we show how humans routinely use available information around them without being aware either that they are doing this or of how unthinkingly they are doing it. With prove-it-to-yourself demonstrations (a favorite method of ours) that engage your eyes and your mind, we demonstrate biases to be unwanted consequences of mental adaptations that may have been evolutionarily helpful in our species’ past.
In “Shades of Truth” (Chapter 2), we lay a basis for understanding why we want to move well beyond long-established research methods that rely on verbal answers to researchers’ questions. We lay out several varieties of “untruths” (lies) that occur in the process of answering questions. These include routine self-deceptions that keep us from recognizing truths about ourselves. Most of the untruths we describe in Chapter 2 are not deliberate attempts to deceive. Rather, they are part of an armory of mental strategies that humans routinely deploy with no active thought about what they are doing.

A crucial moment in the book arrives when we dive “Into the Blindspot” (Chapter 3), where you will have the chance to discover hidden biases of your own. This is where we present and explain the IAT—the invention that made this book possible. At the website implicit.harvard.edu anyone can sample multiple IATs with complete anonymity, adding to the 14 million tests that have already been sampled.

In Chapters 4–7, we explore the psychology of hidden biases by studying how the feelings and beliefs we have toward others are shaped by our feelings and beliefs about the groups to which they belong. As in Chapters 1–3, we use demonstrations that show how mental operations of which we remain unaware can influence judgments and actions in many situations of everyday life.

Chapter 4 takes up how our positive and negative feelings toward other people are influenced by their age, race, ethnicity, sex, social class, and occupation. Chapter 5 describes how the human mind uses these categories to create a huge mental library that rapidly retrieves stereotypical knowledge that we unthinkingly apply to any new person we encounter. Chapter 6 extends this work to describe how stereotypes operate outside of awareness, even when we personally give no credence to those stereotypes and may even reject them. Chapter 7 looks at the distinction of “us” and “them” to understand how we act differently toward in-group members (people who share group membership with us) than toward out-group members, and how this treatment varies depending on whether or not you belong to your society’s dominant group.

By the start of Chapter 8 (“Outsmarting the Machine”) the reader has already encountered many indications that hidden biases can produce unintended damages—interestingly, not only to others but also to oneself. Chapter 8 takes up the question that is central to the goal of preventing the unintended consequences of hidden biases: Are there
effective ways to eliminate hidden biases or, if not, then to neutralize them?

The surge of recent attention to hidden biases was preceded by seven decades of scientific research on non-hidden forms of bias—in other words, prejudice. By far the most intensively scientifically investigated form of prejudice in the U.S. has been (and continues to be) racial prejudice against African Americans. In writing about hidden biases, it was essential for us to take into account what that long history of research has revealed about race prejudice. We do this in two appendices. The first (“Are Americans Racist?”) traces the history of race prejudice in America, in the process addressing a controversial question: Is American racism disappearing, or has it taken on a different, less visible form? The second appendix (“Race, Disadvantage, and Discrimination”) describes what is scientifically known about causes of the lower average status of Black than White Americans. Is race discrimination the cause?

We two met in Columbus, Ohio, in 1980 when Mahzarin arrived from India as a PhD student to work with Tony at Ohio State University. Neither of us is still at Ohio State, but we both remain strongly identified with that institution because we are aware of the role its social psychology PhD program played in our scientific development. The 1980s was a decade in which significant changes were unfolding in our special area of social psychology, called social cognition. Psychology was on the verge of what can now—thirty years later—be recognized as a revolution that introduced new ways of understanding how much of human judgment and behavior are guided by processes that operate outside conscious awareness and conscious control.

The psychological revolution that started in the 1980s was triggered by new methods—ones that could reveal potent mental processes that were not accessible to introspection. This work was happening in studies of how people perceive and remember. We were intrigued by the possibilities represented by the new indirect measures of the mind, and we wondered if those methods could be extended to reveal and explain previously unseen influences on social behavior. Looking back to that period, we can see that we were fortunate to be swept into the vortex of this revolution.

Even though the surge of research on unconscious functioning has not yet peaked, it has already produced significant changes in how human behavior is understood. A quarter
century ago, most psychologists believed that human behavior was primarily guided by conscious thoughts and feelings. Nowadays the majority will readily agree that much of human judgment and behavior is produced with little conscious thought. A quarter century ago, the word “unconscious” was barely to be found in the scientific journals that we read and in which we published our research. Nowadays, the term “unconscious cognition” appears frequently, although it was surpassed in the 1990s by a related term, “implicit cognition.” A quarter century ago, psychologists’ methods for understanding the mind relied mostly on subjects’ verbal reports about their mental states and intentions. Nowadays, research methods are much more diverse, including many that do not at all rely on research subjects’ introspective reports of the contents of their minds or the causes of their behavior.

It was with some trepidation that we referred to the “good people” in this book’s subtitle. Even though—like almost everyone else—we have our opinions about what constitutes goodness in people, we have no special competence (let alone the moral authority) to say who is good and who is not. By “good people” we mean something specific; we refer to all people who intend well and who strive to align their behavior with their good intentions.

We are not at all reluctant to apply in our own lives what we have learned scientifically about the forces now understood to direct our behavior. However, in keeping with our intention not to impose our own values on readers, we have sought to avoid prescriptive statements about how the scientific findings we describe should be put to use in everyday behavior. There is, however, one domain in which it seems entirely proper for our values to guide not only our own behavior but also how we expect others to evaluate the conclusions we reach in this book—our scientific values. Scientific values serve as our guideposts in deciding what research findings warrant presentation as established empirical fact and what theories should be considered valid. A very relevant scientific value in this regard is caution—we need quite a bit of evidence to let empirical findings and scientific theories make the transition from being ideas that we find plausible to being ideas that we are willing to present as established findings or theories.

Another important value for us is skepticism, which includes the willingness to abandon, in the face of opposing evidence, ideas that presently exceed our caution threshold for acceptance. Like all scientists, we do not have the luxury of believing that what now
appears true and valid will always be so. Inevitably, future knowledge will exceed and replace present understanding. If we have done a good job, it may take a few decades for that to happen to the conclusions reached in this book. Among the conclusions that we expect to survive for a while is the idea that hidden-bias blindspots are widespread and that even good people have them.
It is an ordinary day on a college campus. Students and professors of experimental psychology have filed into a lecture hall to listen to a distinguished visiting scientist explain how our minds perceive the physical world. Nothing about his tweed jacket and unkempt hair suggests the challenge he is about to deliver. A few minutes into the lecture, he says matter-of-factly, “As you can plainly see, the two tabletops are exactly the same in shape and size.”

Shuffling in their seats, some in the audience frown while others smile in embarrassment because, as anyone can plainly see, he is dead wrong. Some tilt their heads from side to side, to test if a literal shift in perspective will help. Others wonder whether they should bother staying for the lecture if this nonsense is just the start.

The nonbelievers are caught short, though, when the speaker proceeds to show the truth of his audacious claim. Using an overhead projector, he takes a transparent plastic sheet containing only a single red parallelogram, lays it over the tabletop on the left, and shows that it fits perfectly. He then rotates the plastic sheet clockwise, and places the parallelogram over the tabletop on the right; it fits perfectly there as well. An audible gasp fills the hall as the speaker moves the red frame back and forth, and the room breaks into
laugh. With nothing more than a faint smile the speaker goes on to complete his lecture on how the eye receives, the brain registers, and the mind interprets visual information.

Unconvinced? You can try the test yourself. Find some paper thin enough to trace the outline of one of the tabletops, and then move the outline over to the other tabletop. If you don’t find that the shape of the first tabletop fits identically onto the second tabletop, there can be only one explanation—you’ve botched the tracing job, because the table surfaces are precisely the same. But how can this be?

**VISUAL MINDBUGS**

You, like us, have just succumbed to a famous visual illusion, one that produces an error in the mind’s ability to perceive a pair of objects as they actually are. We will call such errors *mindbugs*—ingrained habits of thought that lead to errors in how we perceive, remember, reason, and make decisions.¹

The psychologist Roger Shepard, a genius who has delighted in the art of confounding, created this illusion called *Turning the Tables*. When we look at the images of the two table surfaces, our retinas do, in fact, receive them as identical in shape and size. In other words, the retina “sees” the tabletops quite accurately. However, when the eye transmits that information to the brain’s visual cortex, where depth is perceived, the trouble begins.

The incorrect perception that the two tabletops are strikingly different in shape occurs effortlessly, because the brain automatically converts the 2-D image that exists both on the page and on the retina into a 3-D interpretation of the tabletop shapes as they must be in the natural world. The automatic processes of the mind, in other words, impose the third dimension of depth onto this scene. And the conscious, reflective processes of the mind accept the illusion unquestioningly. So much so that when encountering the speaker’s assertion that the tabletop outlines are the same, the conscious mind’s first reaction is to consider it to be sheer nonsense.

Natural selection has endowed the minds of humans and other large animals to operate successfully in a three-dimensional world. Having no experience in a world other than a 3-D one, the brain we have continues to perform its conscious perceptual corrections of the tables’ dimensions to make them appear as they would in the traditional 3-D world.²
Contrary to expectation, this error reflects not a weakness of adaptation but rather a triumph, for Shepard’s tabletops highlight the success of a visual system that has adapted effectively to the combination of a two-dimensional retina inside the eye and a three-dimensional world outside. The mind’s automatic understanding of the data is so confident that, as Shepard puts it, “any knowledge or understanding of the illusion we may gain at the intellectual level remains virtually powerless to diminish the magnitude of the illusion.” Take a look at the tables again. The knowledge you now have (that the tables have identical surfaces) has no corrective effect in diminishing the illusion.\(^3\)

Disconcerting as this experience is, it serves as a vivid illustration of a signal property of the mind—it does a great deal of its work automatically, unconsciously, and unintentionally. Mention of the mind’s unconscious operation may summon up for you a visual memory of the bearded, cigar-smoking Sigmund Freud, who rightly gets credit for having brought the term *unconscious* into everyday use. However, an understanding of the unconscious workings of the mind has changed greatly in the century since Freud’s pathbreaking observations. Freud portrayed an omniscient unconscious with complex motives that shape important aspects of human mind and behavior—from dreams to memories to madness, and ultimately to civilization itself. Today, however, Freud’s arguments, detached as they have remained from scientific verification, have a greatly reduced impact on scientific understanding of unconscious mental life.

Instead, the modern conception of the unconscious mind must be credited to another historical figure, one far less known than Freud. A nineteenth-century German physicist and physiologist, Hermann von Helmholtz, offered the name *unbewusster Schluß*, or *unconscious inference*, to describe how an illusion like Shepard’s tabletops might work.\(^4\) Helmholtz aimed to describe the means by which the mind creates from physical data the conscious perceptions that define our ordinary and subjective experiences of “seeing.” Our visual system is capable of being tricked by a simple 2-D image, because an unconscious mental act replaces the 2-D shape of the retinal image with a consciously perceived 3-D shape of the inferred object it suggests.

Now try this: Read the following sixteen words with sufficiently close attention so that you can expect to be able to recognize them when you see them again a few pages hence:

Ant
In the meantime, here’s another striking example of unconscious inference in the form of a checkerboard and cylinder to confound us. When we tell you that the squares marked A and B are exactly the same in their coloring, you will doubtless believe us to be wrong. But take a thick piece of opaque paper, one large enough to cover the entire picture, mark with a point the two squares labeled A and B, and make a circular hole just a bit smaller than the checkerboard square on which each sits. When you look only through the holes and without the rest of the image, you will see that they are indeed identical in color.
Again the culprit is an unconscious inference, a mindbug that automatically goes to work on the image. What causes this remarkable failure of perception? Several features of this checkerboard image are involved, but let us attend to the most obvious ones. First of all, notice that B is surrounded by several dark squares that make it look lighter than it is, merely by contrast; likewise, just the opposite, A is surrounded by adjacent lighter squares that make it seem darker than it actually is. Second, notice the shadow being cast by the cylinder. This darkens the squares within the shadow—including the one marked B—but the mind automatically undoes this darkening to correct for the shadow, lightening our conscious experience of B.

As with the table illusion, the mechanisms that produce this one also exist to enable us to see and understand the world successfully. Ted Adelson, a vision scientist at MIT and creator of this checkershadow image, writes: “As with many so-called illusions, this effect really demonstrates the success rather than the failure of the visual system. The visual system is not very good at being a physical light meter, but that is not its purpose.” Such examples force us to ask a more general question: To what extent has our mind designed wonderfully efficient and accurate methods that fail us miserably when we put them to use in a slightly revised context?
MEMORY MINDBUGS

Think back to the words you memorized earlier, as you examine the list below. As you review each word, without turning back to the original list, try to recall whether each word you see here also appeared in the list you read earlier. If you have paper and pencil handy, and to avoid any doubt about your answers, copy all the words you recall seeing on the previous list and leave out any word that, by your recollection, did not appear before.

Maple Ant Poison Fly Stem Berry Feelers Slimy Birch Wing Leaves Tree Roots Bite Web Bug Small Oak Crawl Acorn Wasp Branch Insect Bee Willow Fright Spider Pine Creepy

You should have left out all twelve tree-related words, starting with maple and ending with pine, for indeed, none of the tree words appeared on the earlier list. You should have also written down all the insect-related words, except one—the word insect itself! That word was not on the original list. If, as is quite likely, you included the word insect, you have demonstrated a powerful but ordinary mindbug that can cause false memories.

In retrospect, it’s easy to see the basis for the false memory for insect. The mind is an automatic association-making machine. When it encounters any information—words, pictures, or even complex ideas—related information automatically comes to mind. In this case, the words in the original list had an insect theme. Unthinkingly, we use that shared theme as we try to remember the past and, in so doing, stumble easily when we come across the word insect itself. Such a memory error is called a false alarm—we mistakenly remember something that did not occur.

In a study conducted at Washington University, 82 percent of the time students remembered seeing words that shared a theme—say, insects—but were not on the original test lists. That huge percentage of errors is especially remarkable when compared to the 75 percent correct memory for words that were actually on the list! In other words, mindbugs can be powerful enough to produce greater recollection of things that didn’t occur than of things that did occur.6

The errors witnessed so far may not seem terribly consequential. What’s the harm, after all, in misremembering a word? But imagine being interrogated about a potential suspect in a crime you have witnessed. Could the false-memory mindbug interfere with your accuracy
in reporting what you saw? If the suspect bears some resemblance to the criminal—for example, has a similar beard—might a false identification result? If so, with what probability?

Elizabeth Loftus is among psychology’s most creative experimentalists. Now at the University of California at Irvine, she has made it her life’s work to study memory mindbugs in eyewitnesses by presenting simulated burglaries, car accidents, and other common mishaps and then testing people’s memories of them. She has found not only that errors in these eyewitness memories are disturbingly frequent but also that even slight changes in the way in which the witness is prompted during questioning to remember an event can alter the content of what is remembered.

In one famous study, Loftus showed witnesses scenes from an automobile accident in which two cars had collided with no personal injury. Later she asked half of the witnesses, “How fast was the car going when it hit the other car?” She asked the other half, “How fast was the car going when it smashed into the other car?” Those who were asked the “smashed” question gave higher estimates of the speed of the vehicle, compared to those who were asked the “hit” question, in addition to which they were more likely to mistakenly insert a memory of broken glass at the accident scene even though there was none in what they saw.7

Psychologists call this mindbug retroactive interference—an influence of after-the-experience information on memory. Loftus gave this a more memorable name: the misinformation effect. Her point is that a small change in language can produce a consequential change in what is remembered, often resulting in mistaken testimony by eyewitnesses who relied on mistaken information.

In recent years it has become clear that the number of wrongful convictions produced by eyewitness errors is substantial.8 From the efforts of the Innocence Project, an organization dedicated to exonerating the wrongfully convicted through DNA testing, 250 people so far have been exonerated by conclusive tests that confirmed their innocence. Of these, 190 cases had been decided based on a mistaken eyewitness account. In other words, in nearly 75 percent of the cases of wrongful conviction, the failure of eyewitness memory (assuming no malign intent on the part of the witness to wrongfully convict) was responsible for tragedies that many societies believe to be so intolerable that their laws explicitly err on
the side of allowing the guilty to walk free.

**AVAILABILITY AND ANCHORING: TWO FAMOUS MINDBUGS**

**Pick the correct answer** in each of the three pairs: Each year, do more people in the United States die from cause (a) or cause (b)?

1. (a) tornados (b) lightning
2. (a) murder (b) suicide
3. (a) car accidents (b) abdominal cancer

Most of us answer (a) in all three cases, when in fact the correct answer to each is (b)—lightning takes more lives than tornados do, suicides are responsible for more deaths than murders are, and abdominal cancer kills more people than do car accidents. Psychologists Daniel Kahneman and Amos Tversky named and described the generic version of this mindbug, calling it the *availability heuristic*. When instances of one type of event (such as death by murder rather than suicide) come more easily to mind than those of another type, we tend to assume that the first event also must occur more frequently in the world. Tornados are more likely to be reported by the media than lightning strikes, perhaps because they affect larger groups of people all at once. This is seemingly reasonable, but it can lead us to overestimate tornados’ death toll. However, greater ease of availability to the mind doesn’t mean greater frequency of occurrence in the world. These kinds of mistakes occur routinely, and are often accompanied with great costs.9

Dan Ariely, a behavioral economist, asked students at MIT to write down the last two digits of their Social Security number on a piece of paper. He then asked them to estimate the price of a keyboard, a trackball, or a design book, items easily familiar to MIT students. Dan collected these two numbers from each person and then computed the correlation between them, looking for a possible relation between the two digits of the Social Security number and the estimated prices. Logically, there is no connection between the two sets of numbers, so the correlation should have been at or close to zero.

In fact, Ariely discovered that there was a substantial correlation between the two sets of numbers. Those for whom the last two digits of their Social Security number happened to
lie between 00 and 19 said they would pay $8.62 on average for the trackball; those with
digits between 20 and 39 were willing to pay more, $11.82; those with digits between 40
and 59 offered up even more, $13.45; and the poor souls whose Social Security numbers
happened to end in digits from 60 to 79 and 80 to 99 offered to pay $21.18 and $26.18—all
for the very same object! ¹⁰

This, the second of the two famous mindbugs, was also discovered by psychologists Daniel
Kahneman and Amos Tversky, who called it anchoring, to capture the idea that the mind
doesn’t search for information in a vacuum. Rather, it starts somewhere, using whatever
information is immediately available as a reference point or “anchor” and then adjusting.
The result, in this case of the random-digit anchor, was the potentially self-harming penalty
of being willing to pay too much. ¹¹

Those who fall prey to the availability and anchoring heuristics are not more feeble-
minded or gullible than others. Each of us is an ever-ready victim. Property values can be
altered by manipulated price anchors that inflate or deflate the actual price. The valuation
of a stock can be influenced more by its suggested market price than its actual value,
perhaps providing some of the explanation for the persistence of financial bubbles. ¹²

SOCIAL MINDBUGS

Human beings are social animals, first and foremost. Other members of our species are
significant to us in ways that little else in the physical world can compete with. For this
reason perhaps, the primate brain has evolved to pay special attention to others of its kind,
and one way in which we do this is to routinely try to predict what might go on in the
minds of others.

Emerging research suggests that selective brain regions appear to be active when we
imagine the thoughts of another person (Does she believe in Christ the Savior?) and when we
try to predict the actions of others (Will he allow my temple to be safe if we build near his
home?). ¹³ These same brain regions do not seem to care when we contemplate the physical
aspects of others, such as their height, weight, or eye color, suggesting that the brain has
evolved specific regions to help with the tasks of social thinking and feeling. That is to say,
other minds matter to us enough that regions of neural real estate are uniquely engaged for
the purpose of making social meaning.
Gordon Bower, a cognitive psychologist at Stanford, was interested in how memory plays a role in important decisions about people. He invited groups of individuals to be members of a jury in a mock trial. The defendant, Mr. Sanders, had run a stop sign while driving, with the consequence of colliding with a garbage truck. Although the defendant’s blood alcohol level was not tested at the time of the accident, he was being tried on the basis of evidence that he might have been drunk while driving. Testimony was presented, with each of two groups of subjects receiving one of the following two descriptions of Mr. Sanders’s behavior at a party just prior to the accident:

(1) On his way out the door, Sanders staggered against a serving table, knocking a bowl to the floor.
(2) On his way out the door, Sanders staggered against a serving table, knocking a bowl of guacamole dip to the floor and splattering guacamole on the white shag carpet.

Should there have been any difference in these two snippets of testimony on judgments of Mr. Sanders’s guilt or innocence? Not at all. The information about the color of the food and its appearance on the white rug were logically irrelevant to his possible drunkenness. But those who heard the testimony with the additional detail about the guacamole were more likely to believe that Sanders was guilty. The culprit here is the vividness with which some events stick in memory. Vividness makes some information more readily available in memory (remember the availability bias?), and what Bower’s experiment shows is that availability through vividness plays a role in skewing a jury’s verdict, presumably with no awareness on the part of jurors that this was happening.

It’s striking that these examples are not about rare or bizarre events that afflict a small group of unusually fragile people. Their ordinariness makes them such compelling phenomena to understand.

It is easy to see why we may give little thought to errors that harm others, such as Mr. Sanders, but not us. Might we be more careful and avoid mindbugs when our own interests are at stake? Imagine an experiment you might do with six friends. Ask a group of three of these friends (randomly chosen from the six) to give three reasons why they love their romantic partners, and ask the other three friends to give nine reasons for the same. Then ask both groups of friends this single question: “How satisfied are you with your
relationship?” Surprisingly, research suggests that you will find that those asked to write only three reasons report greater happiness with their partner and their relationship than those asked to write nine reasons.

The explanation for the bias is counterintuitive but simple: Which of us can easily come up with nine good qualities of a partner? Even canonization requires only two miracles! Those asked to come up with nine reasons have to work harder to come up with the list, and it prompts them to think: “Hmm, this is hard! Is it possible that my partner isn’t as wonderful as I’d imagined?” The experiment that tested this—by Norbert Schwarz at the University of Michigan—found that even important and familiar affections can be manipulated by the availability bias.¹⁴

When judging people’s character, we hardly even recognize just how correct our assessment feels—even when it is based on only a smattering of information. Look at photos of any two strangers and ask yourself:

Which one of these two people seems to be more trustworthy?
Which one will be more competent on the job?
Which one is more likely to dominate others?

It turns out that it is surprisingly easy (we didn’t say accurate) to make such assessments based on nothing more than a static picture. In fact, trying to avoid making such judgments may be far more effortful than making them. The problem, of course, is that these judgments may be not just a little wrong but quite wrong. A face whose features are similar to our own may evoke a feeling of trust when we are deciding to hire a job applicant or choosing a candidate to vote for. And particular facial configurations can lead us all to believe that a person is trustworthy—such as those who have more of a baby face.

Alex Todorov at Princeton University shows that shifting the eyes on a face closer together can make a person look less competent.¹⁵ The implications from such a demonstration should give us pause, enough pause to rethink the manner in which we make important decisions and the degree to which they are even in our own interest. Our interactions with others require not only that we routinely make decisions about them but also that we do so under conditions of less than perfect knowledge. It’s never quite clear
which of two people, Shakeel or Stanley, was responsible for the brawl that erupted in a bar. It’s not easy to say whether Joan or Joe will best lead our team into emerging markets. It’s not easy to say whether Manuel or Mohammed has the necessary skills to serve as a competent Transportation Security Administration checkpoint guard. In dealing with such difficult assessments, we rely on the social group to which the person belongs as the basis for predicting success. Without recognizing it, we automatically pose and answer questions such as: Are people like him trustworthy or not? Is the group she comes from generally smart or dumb? Are people of his kind likely to be violent or peace-loving?

The same mind that viewed the two tables to be different when in fact they were the same is at work here, using membership in social groups rather than the table legs as the contextual cue that generates an unconscious inference. In multiple experiments, we have given people nothing more than a picture of a human face and asked them to tell us whether Mark is likely to enjoy skiing or reading, whether Sally will visit family for the holidays, whether Heather finds shoe shopping to be a nuisance, and so on for dozens of questions. Remarkably, nobody ever responds, “I can’t say!” Using whatever they can eke out from even the most trivial information, people make assessments within a few seconds or fractions of a second, and without any visible discomfort at having to do so. This comfort with judging the likely trustworthiness and competence of others in the absence of any direct information at all—sometimes just a 2-D view of a face—is worth some consideration in itself.

Social mindbugs can give us both false feelings of faith in people we perhaps shouldn’t trust and the opposite—feelings of distrust toward those whom we perhaps should trust. Take as an example Bernard Madoff, perpetrator of the largest investment fraud in U.S. history. While Madoff’s victims were quite ethnically diverse, Jewish philanthropic organizations were particularly numerous among them, suggesting that they were more unwisely trusting of Madoff, with whom they shared a group identity based on religion. A symmetric but opposite outcome involves the evocation of inappropriate distrust that is also in error. As the story goes, when Omar Edwards left his home on the morning of May 28, 2009, he had little reason to predict that he would never return home. But Edwards, a Black police officer, was fatally shot in Harlem when a fellow officer mistook him for a suspect. Such tragic actions tell us that we fail to perceive individuals as individuals. They are often viewed as representatives of particular social groups. Tragedies that arise both from
inappropriate trust and from inappropriate distrust bear the imprint of automatic decisions made on the basis of group membership.

Economists, sociologists, and psychologists have confirmed time and again that the social group to which a person belongs can be isolated as a definitive cause of the disparate treatment he or she receives. Our work has led us to think about single ordinary instances—a smile or a suspicious look, a bank loan approved or rejected, a decision to stop and search, to promote or let go, to investigate with further medical tests or not. Each individual act involves but a single decision that one mind makes about another, and it is here that we must look for mindbugs.

Social mindbugs are not restricted to decisions based on a person’s race or ethnicity. They stem from the existence of psychologically and socially meaningful human groups. Age, gender, religion, class, sexuality, disability, physical attractiveness, profession, and personality are only a few examples, and some are more magnetic than others in drawing us toward them as explanations of behavior. As Jean-Paul Sartre’s famous narration goes, when a woman had difficulty with a particular furrier, she added as an explanation of her troubles with him that he was Jewish. Why, Sartre asked, did she not turn her dislike of the man to be a property of furriers? Why Jews? Mindbugs are at the root of such likes and dislikes, even strong passions. The groups to which people belong seem to be compelling explanations for who they are and what they do and even what they may potentially do, and thereby serve as justification of behavior toward them.

Curiously, social mindbugs affect decisions not only about others but also about ourselves. In a study conducted at Yale’s School of Public Health, Becca Levy showed a stunning correlation—that the negative beliefs about the elderly that elderly people themselves held when they were younger predicted their vulnerability to heart disease when they became older. This result emerged even after controlling for other factors such as depression, smoking, and family history. We take such evidence as suggestive that stereotypes can be harmful not just to the others we assess and evaluate but also to ourselves. In understanding mindbugs, a persuasive reason to take them seriously is self-interest: Stereotypes can negatively affect our actions toward ourselves.

CODA
HERBERT SIMON, a pioneer in studying the mind’s capacities, poked fun at the assumption that people make decisions based on a sensible analysis of the actual value of the options available to them. Human beings, Simon noted, have “neither the facts nor the consistent structure of values nor the reasoning power at their disposal” to make decisions in line with subjective expected utility, which he argued was “a beautiful object deserving a prominent place in Plato’s heaven of ideas ... [but] impossible to employ ... in any literal way in making actual human decisions.”

Give people the option of taking a free $10 bill versus an equally free $20 bill and they will, if rational, take the latter. But as soon as the situation gets even slightly more complex or uncertain, people begin to depart from the economists’ expectation of “utility maximization.” We now know that in our daily choices, whether it is choosing between eating pepperoni or mushroom, visiting Bali or Barcelona, choosing investment banking or carpentry, we do not always act to maximize our own happiness and well-being.

Evidence from the second half of the twentieth century has made it increasingly plausible that human rationality is severely limited. Our task, in this book, is to follow this idea of the bounds on rationality into a particular place—where the questions concern how we judge ourselves, other individuals, and the social collectives to which they and we belong.

As a backdrop to our analysis, we will keep in mind that humanity has a long evolutionary past that has shaped social nature to be what it is today. Our ancestors lived in relatively small, homogeneous groups, surrounded by constant physical danger. In response to the pressures of that environment, they evolved mechanisms that made social choices paramount to mere survival. Several of the mindbugs we have described are consequences of these erstwhile evolutionary triumphs.

But we also recognize that the modern social world in which we live would be unrecognizable to our ancestors. The social contents of our minds, how we think about what’s right and wrong as it involves other people, likely would be incomprehensible to them. Human values and aspirations have changed radically and rapidly, even within the last few generations. Principles of equal rights and fair treatment, values essential to any modern democratic political system, have existed for barely a few centuries. The demands placed on us to survive in the past are not necessarily the same demands that will allow us to thrive now. As just one example, staying away from those who were different or strange
may once have been a safe strategy. Now, it can be financially costly, if the tendency to avoid those who are different keeps a corporation from investing in places distant from the business’s country of origin or from working in a labor market that does not speak their language.

Even a short time ago it would have been unthinkable that we should devise regulations to monitor our behavior—for example, that we should prohibit ourselves from buying a cup of coffee for a federal employee who may be in a position to fund our research. Think about the vast difference in our notion of power, even the seemingly “natural” power of parents over children, when we make laws that dictate that after a certain age children have a right to privacy of information—not to mention changes in our views of torture and the rights of enemy combatants. These changes didn’t happen because a god made it so. They happened because our conscious, reflective minds changed our notion of what is fair and just and what is the right way to conduct ourselves as social beings. We explicitly shifted from ways of being that no doubt were more comfortable and natural to us and imposed on ourselves harder but loftier standards of accountability.

We changed because to the extent that mindbugs slant how we see, remember, reason, and judge, they reveal a particular disparity in us: between our intentions and ideals, on one hand, and our behavior and actions, on the other. The mind sciences have shown that such disparities undermine self-awareness, threaten the ability to consciously control actions, and obfuscate the cherished ideal of self-determination. Understanding how mindbugs erode the coastline of rational thought, and ultimately the very possibility of a just and productive society, requires understanding the mindbugs that are at the root of the disparity between our inner minds and outward actions.