

THE CHARACTERS OF THE STORY

To observe your mind in automatic mode, glance at the image below.



Figure 1

Your experience as you look at the woman's face seamlessly combines what we normally call seeing and intuitive thinking. As surely and quickly as you saw that the young woman's hair is dark, you knew she is angry. Furthermore, what you saw extended into the future. You sensed that this woman is about to say some very unkind words, probably in a loud and strident voice. A premonition of what she was going to do next came to mind automatically and effortlessly. You did not intend to assess her mood or to anticipate what she might do, and your reaction to the picture did not have the

feel of something you did. It just happened to you. It was an instance of fast thinking.

Now look at the following problem:

$$17 \times 24$$

You knew immediately that this is a multiplication problem, and probably knew that you could solve it, with paper and pencil, if not without. You also had some vague intuitive knowledge of the range of possible results: You would be quick to recognize that both 12,609 and 123 are implausible. Without spending some time on the problem, however, you would not be certain that the answer is not 568. A precise solution did not come to mind, and you felt that you could choose whether or not to engage in the computation. If you have not done so yet, you should attempt the multiplication problem now, completing at least part of it.

You experienced slow thinking as you proceeded through a sequence of steps. You first retrieved from memory the cognitive program for multiplication that you learned in school, then you implemented it. Carrying out the computation was a strain. You felt the burden of holding much material in memory, as you needed to keep track of where you were and of where you were going, while holding on to the intermediate result. The process was mental work: deliberate, effortful, and orderly—a prototype of slow thinking. The computation was not only an event in your mind; your body was also involved. Your muscles tensed up, your blood pressure rose, and your heart rate increased. Someone looking closely at your eyes while you tackled this problem would have seen your pupils dilate. Your pupils contracted back to normal size as soon as you ended your work—when you found the answer (which is 408, by the way) or when you gave up.

TWO SYSTEMS

Psychologists have been intensely interested for several decades in the two modes of thinking evoked by the picture of the angry woman and by the multiplication problem, and have offered many labels for them. I adopt terms originally proposed by the psychologists Keith Stanovich and Richard West, and will refer to two systems in the mind, System 1 and System 2.

- *System 1* operates automatically and quickly, with little or no effort and no sense of voluntary control.

- *System 2* allocates attention to the effortful mental activities that demand it, including complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration.

The labels of System 1 and System 2 are widely used in psychology, but I go further than most in this book, which you can read as a psychodrama with two characters.

When we think of ourselves, we identify with System 2, the conscious, reasoning self that has beliefs, makes choices, and decides what to think about and what to do. Although System 2 believes itself to be where the action is, the automatic System 1 is the hero of the book. I describe System 1 as effortlessly originating impressions and feelings that are the main sources of the explicit beliefs and deliberate choices of System 2. The automatic operations of System 1 generate surprisingly complex patterns of ideas, but only the slower System 2 can construct thoughts in an orderly series of steps. I also describe circumstances in which System 2 takes over, overruling the free-wheeling impulses and associations of System 1. You will be invited to think of the two systems as agents with their individual abilities, limitations, and functions.

In rough order of complexity, here are some examples of the automatic activities that are attributed to System 1:

- Detect that one object is more distant than another.
- Orient to the source of a sudden sound.
- Complete the phrase “bread and . . .”
- Make a “disgust face” when shown a horrible picture.
- Detect hostility in a voice.
- Answer to $2 + 2 = ?$
- Read words on large billboards.
- Drive a car on an empty road.
- Find a strong move in chess (if you are a chess master).
- Understand simple sentences.
- Recognize that a “meek and tidy soul with a passion for detail” resembles an occupational stereotype.

All these mental events belong with the angry woman—they occur automatically and require little or no effort. The capabilities of System 1 include innate skills that we share with other animals. We are born prepared to perceive the world around us, recognize objects, orient attention, avoid losses,

and fear spiders. Other mental activities become fast and automatic through prolonged practice. System 1 has learned associations between ideas (the capital of France?); it has also learned skills such as reading and understanding nuances of social situations. Some skills, such as finding strong chess moves, are acquired only by specialized experts. Others are widely shared. Detecting the similarity of a personality sketch to an occupational stereotype requires broad knowledge of the language and the culture, which most of us possess. The knowledge is stored in memory and accessed without intention and without effort.

Several of the mental actions in the list are completely involuntary. You cannot refrain from understanding simple sentences in your own language or from orienting to a loud unexpected sound, nor can you prevent yourself from knowing that $2 + 2 = 4$ or from thinking of Paris when the capital of France is mentioned. Other activities, such as chewing, are susceptible to voluntary control but normally run on automatic pilot. The control of attention is shared by the two systems. Orienting to a loud sound is normally an involuntary operation of System 1, which immediately mobilizes the voluntary attention of System 2. You may be able to resist turning toward the source of a loud and offensive comment at a crowded party, but even if your head does not move, your attention is initially directed to it, at least for a while. However, attention can be moved away from an unwanted focus, primarily by focusing intently on another target.

The highly diverse operations of System 2 have one feature in common: they require attention and are disrupted when attention is drawn away. Here are some examples:

- Brace for the starter gun in a race.
- Focus attention on the clowns in the circus.
- Focus on the voice of a particular person in a crowded and noisy room.
- Look for a woman with white hair.
- Search memory to identify a surprising sound.
- Maintain a faster walking speed than is natural for you.
- Monitor the appropriateness of your behavior in a social situation.
- Count the occurrences of the letter *a* in a page of text.
- Tell someone your phone number.
- Park in a narrow space (for most people except garage attendants).
- Compare two washing machines for overall value.
- Fill out a tax form.
- Check the validity of a complex logical argument.

In all these situations you must pay attention, and you will perform less well, or not at all, if you are not ready or if your attention is directed inappropriately. System 2 has some ability to change the way System 1 works, by programming the normally automatic functions of attention and memory. When waiting for a relative at a busy train station, for example, you can set yourself at will to look for a white-haired woman or a bearded man, and thereby increase the likelihood of detecting your relative from a distance. You can set your memory to search for capital cities that start with *N* or for French existentialist novels. And when you rent a car at London's Heathrow Airport, the attendant will probably remind you that "we drive on the left side of the road over here." In all these cases, you are asked to do something that does not come naturally, and you will find that the consistent maintenance of a set requires continuous exertion of at least some effort.

The often-used phrase "pay attention" is apt: you dispose of a limited budget of attention that you can allocate to activities, and if you try to go beyond your budget, you will fail. It is the mark of effortful activities that they interfere with each other, which is why it is difficult or impossible to conduct several at once. You could not compute the product of 17×24 while making a left turn into dense traffic, and you certainly should not try. You can do several things at once, but only if they are easy and un-demanding. You are probably safe carrying on a conversation with a passenger while driving on an empty highway, and many parents have discovered, perhaps with some guilt, that they can read a story to a child while thinking of something else.

Everyone has some awareness of the limited capacity of attention, and our social behavior makes allowances for these limitations. When the driver of a car is overtaking a truck on a narrow road, for example, adult passengers quite sensibly stop talking. They know that distracting the driver is not a good idea, and they also suspect that he is temporarily deaf and will not hear what they say.

Intense focusing on a task can make people effectively blind, even to stimuli that normally attract attention. The most dramatic demonstration was offered by Christopher Chabris and Daniel Simons in their book *The Invisible Gorilla*. They constructed a short film of two teams passing basketballs, one team wearing white shirts, the other wearing black. The viewers of the film are instructed to count the number of passes made by the white team, ignoring the black players. This task is difficult and completely absorbing. Halfway through the video, a woman wearing a gorilla suit appears, crosses the court, thumps her chest, and moves on. The gorilla is in view for

9 seconds. Many thousands of people have seen the video, and about half of them do not notice anything unusual. It is the counting task—and especially the instruction to ignore one of the teams—that causes the blindness. No one who watches the video without that task would miss the gorilla. Seeing and orienting are automatic functions of System 1, but they depend on the allocation of some attention to the relevant stimulus. The authors note that the most remarkable observation of their study is that people find its results very surprising. Indeed, the viewers who fail to see the gorilla are initially sure that it was not there—they cannot imagine missing such a striking event. The gorilla study illustrates two important facts about our minds: we can be blind to the obvious, and we are also blind to our blindness.

PLOT SYNOPSIS

The interaction of the two systems is a recurrent theme of the book, and a brief synopsis of the plot is in order. In the story I will tell, Systems 1 and 2 are both active whenever we are awake. System 1 runs automatically and System 2 is normally in a comfortable low-effort mode, in which only a fraction of its capacity is engaged. System 1 continuously generates suggestions for System 2: impressions, intuitions, intentions, and feelings. If endorsed by System 2, impressions and intuitions turn into beliefs, and impulses turn into voluntary actions. When all goes smoothly, which is most of the time, System 2 adopts the suggestions of System 1 with little or no modification. You generally believe your impressions and act on your desires, and that is fine—usually.

When System 1 runs into difficulty, it calls on System 2 to support more detailed and specific processing that may solve the problem of the moment. System 2 is mobilized when a question arises for which System 1 does not offer an answer, as probably happened to you when you encountered the multiplication problem 17×24 . You can also feel a surge of conscious attention whenever you are surprised. System 2 is activated when an event is detected that violates the model of the world that System 1 maintains. In that world, lamps do not jump, cats do not bark, and gorillas do not cross basketball courts. The gorilla experiment demonstrates that some attention is needed for the surprising stimulus to be detected. Surprise then activates and orients your attention: you will stare, and you will search your memory for a story that makes sense of the surprising event. System 2 is also credited with the continuous monitoring of your own behavior—the control that keeps you polite when you are angry, and alert when you are driving at night.

System 2 is mobilized to increased effort when it detects an error about to be made. Remember a time when you almost blurted out an offensive remark and note how hard you worked to restore control. In summary, most of what you (your System 2) think and do originates in your System 1, but System 2 takes over when things get difficult, and it normally has the last word.

The division of labor between System 1 and System 2 is highly efficient: it minimizes effort and optimizes performance. The arrangement works well most of the time because System 1 is generally very good at what it does: its models of familiar situations are accurate, its short-term predictions are usually accurate as well, and its initial reactions to challenges are swift and generally appropriate. System 1 has biases, however, systematic errors that it is prone to make in specified circumstances. As we shall see, it sometimes answers easier questions than the one it was asked, and it has little understanding of logic and statistics. One further limitation of System 1 is that it cannot be turned off. If you are shown a word on the screen in a language you know, you will read it—unless your attention is totally focused elsewhere.

CONFLICT

Figure 2 is a variant of a classic experiment that produces a conflict between the two systems. You should try the exercise before reading on.

Your first task is to go down both columns, calling out whether each word is printed in lowercase or in uppercase. When you are done with the first task, go down both columns again, saying whether each word is printed to the left or to the right of center by saying (or whispering to yourself) "LEFT" or "RIGHT."

LEFT	upper	lower	LOWER
right			
RIGHT	upper	UPPER	lower
			LOWER
LEFT	right	upper	upper

Figure 2

You were almost certainly successful in saying the correct words in both tasks, and you surely discovered that some parts of each task were much easier than others. When you identified upper- and lowercase, the left-hand column was easy and the right-hand column caused you to slow down and perhaps to stammer or stumble. When you named the position of words, the left-hand column was difficult and the right-hand column was much easier.

"These tasks engage System 2, because saying "upper/lower" or "right/left" is not what you routinely do when looking down a column of words. One of the things you did to set yourself for the task was to program your memory so that the relevant words (*upper* and *lower* for the first task) were "on the tip of your tongue." The prioritizing of the chosen words is effective and the mild temptation to read other words was fairly easy to resist when you went through the first column. But the second column was different, because it contained words for which you were set, and you could not ignore them. You were mostly able to respond correctly, but overcoming the competing response was a strain, and it slowed you down. You experienced a conflict between a task that you intended to carry out and an automatic response that interfered with it.

Conflict between an automatic reaction and an intention to control it is common in our lives. We are all familiar with the experience of trying not to stare at the oddly dressed couple at the neighboring table in a restaurant. We also know what it is like to force our attention on a boring book, when we constantly find ourselves returning to the point at which the reading lost its meaning. Where winters are hard, many drivers have memories of their car skidding out of control on the ice and of the struggle to follow well-rehearsed instructions that negate what they would naturally do: "Steer into the skid, and whatever you do, do not touch the brakes!" And every human being has had the experience of *not* telling someone to go to hell. One of the tasks of System 2 is to overcome the impulses of System 1. In other words, System 2 is in charge of self-control.

ILLUSIONS

To appreciate the autonomy of System 1, as well as the distinction between impressions and beliefs, take a good look at figure 3.

This picture is unremarkable: two horizontal lines of different lengths, with fins appended, pointing in different directions. The bottom line is obviously longer than the one above it. That is what we all see, and we

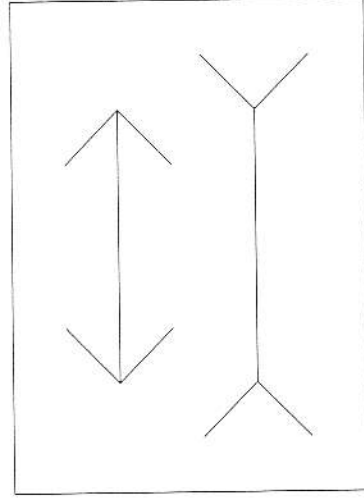


Figure 3

naturally believe what we see. If you have already encountered this image, however, you recognize it as the famous Müller-Lyer illusion. As you can easily confirm by measuring them with a ruler, the horizontal lines are in fact identical in length.

Now that you have measured the lines, you—your System 2, the conscious being you call "I"—have a new belief: you *know* that the lines are equally long. If asked about their length, you will say what you know. But you still *see* the bottom line as longer. You have chosen to believe the measurement, but you cannot prevent System 1 from doing its thing; you cannot decide to see the lines as equal, although you know they are. To resist the illusion, there is only one thing you can do: you must learn to mistrust your impressions of the length of lines when fins are attached to them. To implement that rule, you must be able to recognize the illusory pattern and recall what you know about it. If you can do this, you will never again be fooled by the Müller-Lyer illusion. But you will still see one line as longer than the other.

Not all illusions are visual. There are illusions of thought, which we call *cognitive illusions*. As a graduate student, I attended some courses on the art and science of psychotherapy. During one of these lectures, our teacher imparted a morsel of clinical wisdom. This is what he told us: "You will from time to time meet a patient who shares a disturbing tale of multiple mistakes in his previous treatment. He has been seen by several clinicians, and all failed him. The patient can lucidly describe how his therapists misunderstood him, but he has quickly perceived that you are different. You share the same feeling, are convinced that you understand him, and will be

able to help." At this point my teacher raised his voice as he said, "Do not even *think* of taking on this patient! Throw him out of the office! He is most likely a psychopath and you will not be able to help him."

Many years later I learned that the teacher had warned us against psychopathic charm, and the leading authority in the study of psychopathy confirmed that the teacher's advice was sound. The analogy to the Müller-Lyer illusion is close. What we were being taught was not how to feel about that patient. Our teacher took it for granted that the sympathy we would feel for the patient would not be under our control; it would arise from System 1. Furthermore, we were not being taught to be generally suspicious of our feelings about patients. We were told that a strong attraction to a patient with a repeated history of failed treatment is a danger sign—like the fins on the parallel lines. It is an illusion—a cognitive illusion—and I (System 2) was taught how to recognize it and advised not to believe it or act on it.

The question that is most often asked about cognitive illusions is whether they can be overcome. The message of these examples is not encouraging. Because System 1 operates automatically and cannot be turned off at will, errors of intuitive thought are often difficult to prevent. Biases cannot always be avoided, because System 2 may have no clue to the error. Even when cues to likely errors are available, errors can be prevented only by the enhanced monitoring and effortful activity of System 2. As a way to live your life, however, continuous vigilance is not necessarily good, and it is certainly impractical. Constantly questioning our own thinking would be impossibly tedious, and System 2 is much too slow and inefficient to serve as a substitute for System 1 in making routine decisions. The best we can do is a compromise: learn to recognize situations in which mistakes are likely and try harder to avoid significant mistakes when the stakes are high. The premise of this book is that it is easier to recognize other people's mistakes than our own.

USEFUL FICTIONS

You have been invited to think of the two systems as agents within the mind, with their individual personalities, abilities, and limitations. I will often use sentences in which the systems are the subjects, such as, "System 2 calculates products."

The use of such language is considered a sin in the professional circles in which I travel, because it seems to explain the thoughts and actions of a

person by the thoughts and actions of little people inside the person's head. Grammatically the sentence about System 2 is similar to "The butler steals the petty cash." My colleagues would point out that the butler's action actually explains the disappearance of the cash, and they rightly question whether the sentence about System 2 explains how products are calculated. My answer is that the brief active sentence that attributes calculation to System 2 is intended as a description, not an explanation. It is meaningful only because of what you already know about System 2. It is shorthand for the following: "Mental arithmetic is a voluntary activity that requires effort, should not be performed while making a left turn, and is associated with dilated pupils and an accelerated heart rate."

Similarly, the statement that "highway driving under routine conditions is left to System 1" means that steering the car around a bend is automatic and almost effortless. It also implies that an experienced driver can drive on an empty highway while conducting a conversation. Finally, "System 2 prevented James from reacting foolishly to the insult" means that James would have been more aggressive in his response if his capacity for effortful control had been disrupted (for example, if he had been drunk).

System 1 and System 2 are so central to the story I tell in this book that I must make it absolutely clear that they are fictitious characters. Systems 1 and 2 are not systems in the standard sense of entities with interacting aspects or parts. And there is no one part of the brain that either of the systems would call home. You may well ask: What is the point of introducing fictitious characters with ugly names into a serious book? The answer is that the characters are useful because of some quirks of our minds, yours and mine. A sentence is understood more easily if it describes what an agent (System 2) does than if it describes what something is, what properties it has. In other words, "System 2" is a better subject for a sentence than "mental arithmetic." The mind—especially System 1—appears to have a special aptitude for the construction and interpretation of stories about active agents, who have personalities, habits, and abilities. You quickly formed a bad opinion of the thieving butler, you expect more bad behavior from him, and you will remember him for a while. This is also my hope for the language of systems.

Why call them System 1 and System 2 rather than the more descriptive "automatic system" and "effortful system"? The reason is simple: "Automatic system" takes longer to say than "System 1" and therefore takes more space

in your working memory. This matters, because anything that occupies your working memory reduces your ability to think. You should treat "System 1" and "System 2" as nicknames, like Bob and Joe, identifying characters that you will get to know over the course of this book. The fictitious systems make it easier for me to think about judgment and choice, and will make it easier for you to understand what I say.

SPEAKING OF SYSTEM 1 AND SYSTEM 2

"He had an impression, but some of his impressions are illusions."

"This was a pure System 1 response. She reacted to the threat before she recognized it."

"This is your System 1 talking. Slow down and let your System 2 take control."

2

ATTENTION AND EFFORT

In the unlikely event of this book being made into a film, System 2 would be a supporting character who believes herself to be the hero. The defining feature of System 2, in this story, is that its operations are effortful, and one of its main characteristics is laziness, a reluctance to invest more effort than is strictly necessary. As a consequence, the thoughts and actions that System 2 believes it has chosen are often guided by the figure at the center of the story, System 1. However, there are vital tasks that only System 2 can perform because they require effort and acts of self-control in which the intuitions and impulses of System 1 are overcome.

MENTAL EFFORT

If you wish to experience your System 2 working at full tilt, the following exercise will do; it should bring you to the limits of your cognitive abilities within 5 seconds. To start, make up several strings of 4 digits, all different, and write each string on an index card. Place a blank card on top of the deck. The task that you will perform is called Add-1. Here is how it goes:

Start beating a steady rhythm (or better yet, set a metronome at 1/sec).

Remove the blank card and read the four digits aloud. Wait for two beats, then report a string in which each of the original digits is incremented by 1.

If the digits on the card are 5294, the correct response is 6305. Keeping the rhythm is important.

Few people can cope with more than four digits in the Add-1 task, but if you want a harder challenge, please try Add-3.

If you would like to know what your body is doing while your mind is hard at work, set up two piles of books on a sturdy table, place a video camera on one and lean your chin on the other, get the video going, and stare at the camera lens while you work on Add-1 or Add-3 exercises. Later, you will find in the changing size of your pupils a faithful record of how hard you worked.

I have a long personal history with the Add-1 task. Early in my career I spent a year at the University of Michigan, as a visitor in a laboratory that studied hypnosis. Casting about for a useful topic of research, I found an article in *Scientific American* in which the psychologist Eckhard Hess described the pupil of the eye as a window to the soul. I reread it recently and again found it inspiring. It begins with Hess reporting that his wife had noticed his pupils widening as he watched beautiful nature pictures, and it ends with two striking pictures of the same good-looking woman, who somehow appears much more attractive in one than in the other. There is only one difference: the pupils of the eyes appear dilated in the attractive picture and constricted in the other. Hess also wrote of belladonna, a pupil-dilating substance that was used as a cosmetic, and of bazaar shoppers who wear dark glasses in order to hide their level of interest from merchants.

One of Hess's findings especially captured my attention. He had noticed that the pupils are sensitive indicators of mental effort—they dilate substantially when people multiply two-digit numbers, and they dilate more if the problems are hard than if they are easy. His observations indicated that the response to mental effort is distinct from emotional arousal. Hess's work did not have much to do with hypnosis, but I concluded that the idea of a visible indication of mental effort had promise as a research topic. A graduate student in the lab, Jackson Beatty, shared my enthusiasm and we got to work.

Beatty and I developed a setup similar to an optician's examination room, in which the experimental participant leaned her head on a chin-and-forehead rest and stared at a camera while listening to prerecorded information and answering questions on the recorded beats of a metronome. The beats triggered an infrared flash every second, causing a picture to be taken. At the end of each experimental session, we would rush to have the film developed, project the images of the pupil on a screen, and go to work with a ruler. The method was a perfect fit for young and impatient

researchers: we knew our results almost immediately, and they always told a clear story.

Beatty and I focused on paced tasks, such as Add-1, in which we knew precisely what was on the subject's mind at any time. We recorded strings of digits on beats of the metronome and instructed the subject to repeat or transform the digits one by one, maintaining the same rhythm. We soon discovered that the size of the pupil varied second by second, reflecting the changing demands of the task. The shape of the response was an inverted V. As you experienced it if you tried Add-1 or Add-3, effort builds up with every added digit that you hear, reaches an almost intolerable peak as you rush to produce a transformed string during and immediately after the pause, and relaxes gradually as you "unload" your short-term memory. The pupil data corresponded precisely to subjective experience: longer strings reliably caused larger dilations, the transformation task compounded the effort, and the peak of pupil size coincided with maximum effort. Add-1 with four digits caused a larger dilation than the task of holding seven digits for immediate recall. Add-3, which is much more difficult, is the most demanding that I ever observed. In the first 5 seconds, the pupil dilates by about 50% of its original area and heart rate increases by about 7 beats per minute. This is as hard as people can work—they give up if more is asked of them. When we exposed our subjects to more digits than they could remember, their pupils stopped dilating or actually shrank.

We worked for some months in a spacious basement suite in which we had set up a closed-circuit system that projected an image of the subject's pupil on a screen in the corridor; we also could hear what was happening in the laboratory. The diameter of the projected pupil was about a foot; watching it dilate and contract when the participant was at work was a fascinating sight, quite an attraction for visitors in our lab. We amused ourselves and impressed our guests by our ability to divine when the participant gave up on a task. During a mental multiplication, the pupil normally dilated to a large size within a few seconds and stayed large as long as the individual kept working on the problem; it contracted immediately when she found a solution or gave up. As we watched from the corridor, we would sometimes surprise both the owner of the pupil and our guests by asking, "Why did you stop working just now?" The answer from inside the lab was often, "How did you know?" to which we would reply, "We have a window to your soul."

The casual observations we made from the corridor were sometimes as informative as the formal experiments. I made a significant discovery as I

was idly watching a woman's pupil during a break between two tasks. She had kept her position on the chin rest, so I could see the image of her eye while she engaged in routine conversation with the experimenter. I was surprised to see that the pupil remained small and did not noticeably dilate as she talked and listened. Unlike the tasks that we were studying, the mundane conversation apparently demanded little or no effort—no more than retaining two or three digits. This was a eureka moment: I realized that the tasks we had chosen for study were exceptionally effortful. An image came to mind: mental life—today I would speak of the life of System 2—is normally conducted at the pace of a comfortable walk, sometimes interrupted by episodes of jogging and on rare occasions by a frantic sprint. The Add-1 and Add-3 exercises are sprints, and casual chatting is a stroll.

We found that people, when engaged in a mental sprint, may become effectively blind. The authors of *The Invisible Gorilla* had made the gorilla “invisible” by keeping the observers intensely busy counting passes. We reported a rather less dramatic example of blindness during Add-1. Our subjects were exposed to a series of rapidly flashing letters while they worked. They were told to give the task complete priority, but they were also asked to report, at the end of the digit task, whether the letter *K* had appeared at any time during the trial. The main finding was that the ability to detect and report the target letter changed in the course of the 10 seconds of the exercise. The observers almost never missed a *K* that was shown at the beginning or near the end of the Add-1 task but they missed the target almost half the time when mental effort was at its peak, although we had pictures of their wide-open eye staring straight at it. Failures of detection followed the same inverted-*V* pattern as the dilating pupil. The similarity was reassuring: the pupil was a good measure of the physical arousal that accompanies mental effort, and we could go ahead and use it to understand how the mind works.

Much like the electricity meter outside your house or apartment, the pupils offer an index of the current rate at which mental energy is used. The analogy goes deep. Your use of electricity depends on what you choose to do, whether to light a room or toast a piece of bread. When you turn on a bulb or a toaster, it draws the energy it needs but no more. Similarly, we decide what to do, but we have limited control over the effort of doing it. Suppose you are shown four digits, say, 9462, and told that your life depends on holding them in memory for 10 seconds. However much you want to live, you cannot exert as much effort in this task as you would be forced to invest to complete an Add-3 transformation on the same digits.

System 2 and the electrical circuits in your home both have limited capacity, but they respond differently to threatened overload. A breaker trips when the demand for current is excessive, causing all devices on that circuit to lose power at once. In contrast, the response to mental overload is selective and precise: System 2 protects the most important activity, so it receives the attention it needs; “spare capacity” is allocated second by second to other tasks. In our version of the gorilla experiment, we instructed the participants to assign priority to the digit task. We know that they followed that instruction, because the timing of the visual target had no effect on the main task. If the critical letter was presented at a time of high demand, the subjects simply did not see it. When the transformation task was less demanding, detection performance was better.

The sophisticated allocation of attention has been honed by a long evolutionary history. Orienting and responding quickly to the gravest threats or most promising opportunities improved the chance of survival, and this capability is certainly not restricted to humans. Even in modern humans, System 1 takes over in emergencies and assigns total priority to self-protective actions. Imagine yourself at the wheel of a car that unexpectedly skids on a large oil slick. You will find that you have responded to the threat before you became fully conscious of it.

Beatty and I worked together for only a year, but our collaboration had a large effect on our subsequent careers. He eventually became the leading authority on “cognitive pupillometry,” and I wrote a book titled *Attention and Effort*, which was based in large part on what we learned together and on follow-up research I did at Harvard the following year. We learned a great deal about the working mind—which I now think of as System 2—from measuring pupils in a wide variety of tasks.

As you become skilled in a task, its demand for energy diminishes. Studies of the brain have shown that the pattern of activity associated with an action changes as skill increases, with fewer brain regions involved. Talent has similar effects. Highly intelligent individuals need less effort to solve the same problems, as indicated by both pupil size and brain activity. A general “law of least effort” applies to cognitive as well as physical exertion. The law asserts that if there are several ways of achieving the same goal, people will eventually gravitate to the least demanding course of action. In the economy of action, effort is a cost, and the acquisition of skill is driven by the balance of benefits and costs. Laziness is built deep into our nature.

The tasks that we studied varied considerably in their effects on the pupil. At baseline, our subjects were awake, aware, and ready to engage in a

task—probably at a higher level of arousal and cognitive readiness than usual. Holding one or two digits in memory or learning to associate a word with a digit (3 = door) produced reliable effects on momentary arousal above that baseline, but the effects were minuscule, only 5% of the increase in pupil diameter associated with Add-3. A task that required discriminating between the pitch of two tones yielded significantly larger dilations. Recent research has shown that inhibiting the tendency to read distracting words (as in figure 2 of the preceding chapter) also induces moderate effort. Tests of short-term memory for six or seven digits were more effortful. As you can experience, the request to retrieve and say aloud your phone number or your spouse's birthday also requires a brief but significant effort, because the entire string must be held in memory as a response is organized. Mental multiplication of two-digit numbers and the Add-3 task are near the limit of what most people can do.

What makes some cognitive operations more demanding and effortful than others? What outcomes must we purchase in the currency of attention? What can System 2 do that System 1 cannot? We now have tentative answers to these questions.

Effort is required to maintain simultaneously in memory several ideas that require separate actions, or that need to be combined according to a rule—rehearsing your shopping list as you enter the supermarket, choosing between the fish and the veal at a restaurant, or combining a surprising result from a survey with the information that the sample was small, for example. System 2 is the only one that can follow rules, compare objects on several attributes, and make deliberate choices between options. The automatic System 1 does not have these capabilities. System 1 detects simple relations (“they are all alike,” “the son is much taller than the father”) and excels at integrating information about one thing, but it does not deal with multiple distinct topics at once, nor is it adept at using purely statistical information. System 1 will detect that a person described as “a meek and tidy soul, with a need for order and structure, and a passion for detail” resembles a caricature librarian, but combining this intuition with knowledge about the small number of librarians is a task that only System 2 can perform—if System 2 knows how to do so, which is true of few people.

A crucial capability of System 2 is the adoption of “task sets”: it can program memory to obey an instruction that overrides habitual responses. Consider the following: Count all occurrences of the letter *f* in this page. This is not a task you have ever performed before and it will not come naturally to you, but your System 2 can take it on. It will be effortful to set

yourself up for this exercise, and effortful to carry it out, though you will surely improve with practice. Psychologists speak of “executive control” to describe the adoption and termination of task sets, and neuroscientists have identified the main regions of the brain that serve the executive function. One of these regions is involved whenever a conflict must be resolved. Another is the prefrontal area of the brain, a region that is substantially more developed in humans than in other primates, and is involved in operations that we associate with intelligence.

Now suppose that at the end of the page you get another instruction: count all the commas in the next page. This will be harder, because you will have to overcome the newly acquired tendency to focus attention on the letter *f*. One of the significant discoveries of cognitive psychologists in recent decades is that switching from one task to another is effortful, especially under time pressure. The need for rapid switching is one of the reasons that Add-3 and mental multiplication are so difficult. To perform the Add-3 task, you must hold several digits in your working memory at the same time, associating each with a particular operation: some digits are in the queue to be transformed, one is in the process of transformation, and others, already transformed, are retained for reporting. Modern tests of working memory require the individual to switch repeatedly between two demanding tasks, retaining the results of one operation while performing the other. People who do well on these tests tend to do well on tests of general intelligence. However, the ability to control attention is not simply a measure of intelligence; measures of efficiency in the control of attention predict performance of air traffic controllers and of Israeli Air Force pilots beyond the effects of intelligence.

Time pressure is another driver of effort. As you carried out the Add-3 exercise, the rush was imposed in part by the metronome and in part by the load on memory. Like a juggler with several balls in the air, you cannot afford to slow down; the rate at which material decays in memory forces the pace, driving you to refresh and rehearse information before it is lost. Any task that requires you to keep several ideas in mind at the same time has the same hurried character. Unless you have the good fortune of a capacious working memory, you may be forced to work uncomfortably hard. The most effortful forms of slow thinking are those that require you to think fast.

You surely observed as you performed Add-3 how unusual it is for your mind to work so hard. Even if you think for a living, few of the mental tasks in which you engage in the course of a working day are as demanding as

Add-3, or even as demanding as storing six digits for immediate recall. We normally avoid mental overload by dividing our tasks into multiple easy steps, committing intermediate results to long-term memory or to paper rather than to an easily overloaded working memory. We cover long distances by taking our time and conduct our mental lives by the law of least effort.

SPEAKING OF ATTENTION AND EFFORT

"I won't try to solve this while driving. This is a pupil-dilating task. It requires mental effort!"

"The law of least effort is operating here. He will think as little as possible."

"She did not forget about the meeting. She was completely focused on something else when the meeting was set and she just didn't hear you."

"What came quickly to my mind was an intuition from System 1. I'll have to start over and search my memory deliberately."

3

THE LAZY CONTROLLER

I spend a few months each year in Berkeley, and one of my great pleasures there is a daily four-mile walk on a marked path in the hills, with a fine view of San Francisco Bay. I usually keep track of my time and have learned a fair amount about effort from doing so. I have found a speed, about 17 minutes for a mile, which I experience as a stroll. I certainly exert physical effort and burn more calories at that speed than if I sat in a recliner, but I experience no strain, no conflict, and no need to push myself. I am also able to think and work while walking at that rate. Indeed, I suspect that the mild physical arousal of the walk may spill over into greater mental alertness.

System 2 also has a natural speed. You expend some mental energy in random thoughts and in monitoring what goes on around you even when your mind does nothing in particular, but there is little strain. Unless you are in a situation that makes you unusually wary or self-conscious, monitoring what happens in the environment or inside your head demands little effort. You make many small decisions as you drive your car, absorb some information as you read the newspaper, and conduct routine exchanges of pleasantries with a spouse or a colleague, all with little effort and no strain. Just like a stroll.

It is normally easy and actually quite pleasant to walk and think at the same time, but at the extremes these activities appear to compete for the limited resources of System 2. You can confirm this claim by a simple experiment. While walking comfortably with a friend, ask him to compute 23×78 in his head, and to do so immediately. He will almost certainly stop

in his tracks. My experience is that I can think while strolling but cannot engage in mental work that imposes a heavy load on short-term memory. If I must construct an intricate argument under time pressure, I would rather be still, and I would prefer sitting to standing. Of course, not all slow thinking requires that form of intense concentration and effortful computation—I did the best thinking of my life on leisurely walks with Amos.

Accelerating beyond my strolling speed completely changes the experience of walking, because the transition to a faster walk brings about a sharp deterioration in my ability to think coherently. As I speed up, my attention is drawn with increasing frequency to the experience of walking and to the deliberate maintenance of the faster pace. My ability to bring a train of thought to a conclusion is impaired accordingly. At the highest speed I can sustain on the hills, about 14 minutes for a mile, I do not even try to think of anything else. In addition to the physical effort of moving my body rapidly along the path, a mental effort of self-control is needed to resist the urge to slow down. Self-control and deliberate thought apparently draw on the same limited budget of effort.

For most of us, most of the time, the maintenance of a coherent train of thought and the occasional engagement in effortful thinking also require self-control. Although I have not conducted a systematic survey, I suspect that frequent switching of tasks and speeded-up mental work are not intrinsically pleasurable, and that people avoid them when possible. This is how the law of least effort comes to be a law. Even in the absence of time pressure, maintaining a coherent train of thought requires discipline. An observer of the number of times I look at e-mail or investigate the refrigerator during an hour of writing could reasonably infer an urge to escape and conclude that keeping at it requires more self-control than I can readily muster.

Fortunately, cognitive work is not always aversive, and people sometimes expend considerable effort for long periods of time without having to exert willpower. The psychologist Mihaly Csikszentmihalyi (pronounced six-cent-mihaly) has done more than anyone else to study this state of effortless attending, and the name he proposed for it, *flow*, has become part of the language. People who experience flow describe it as “a state of effortless concentration so deep that they lose their sense of time, of themselves, of their problems,” and their descriptions of the joy of that state are so compelling that Csikszentmihalyi has called it an “optimal experience.” Many activities can induce a sense of flow, from painting to racing motorcycles—and for some fortunate authors I know, even writing a book is often an op-

timal experience. Flow neatly separates the two forms of effort: concentration on the task and the deliberate control of attention. Riding a motorcycle at 150 miles an hour and playing a competitive game of chess are certainly very effortful. In a state of flow, however, maintaining focused attention on these absorbing activities requires no exertion of self-control, thereby freeing resources to be directed to the task at hand.

THE BUSY AND DEPLETED SYSTEM 2

It is now a well-established proposition that both self-control and cognitive effort are forms of mental work. Several psychological studies have shown that people who are simultaneously challenged by a demanding cognitive task and by a temptation are more likely to yield to the temptation. Imagine that you are asked to retain a list of seven digits for a minute or two. You are told that remembering the digits is your top priority. While your attention is focused on the digits, you are offered a choice between two desserts: a sinful chocolate cake and a virtuous fruit salad. The evidence suggests that you would be more likely to select the tempting chocolate cake when your mind is loaded with digits. System 1 has more influence on behavior when System 2 is busy, and it has a sweet tooth.

People who are *cognitively busy* are also more likely to make selfish choices, use sexist language, and make superficial judgments in social situations. Memorizing and repeating digits loosens the hold of System 2 on behavior, but of course cognitive load is not the only cause of weakened self-control. A few drinks have the same effect, as does a sleepless night. The self-control of morning people is impaired at night; the reverse is true of night people. Too much concern about how well one is doing in a task sometimes disrupts performance by loading short-term memory with pointless anxious thoughts. The conclusion is straightforward: self-control requires attention and effort. Another way of saying this is that controlling thoughts and behaviors is one of the tasks that System 2 performs.

A series of surprising experiments by the psychologist Roy Baumeister and his colleagues has shown conclusively that all variants of voluntary effort—cognitive, emotional, or physical—draw at least partly on a shared pool of mental energy. Their experiments involve successive rather than simultaneous tasks.

Baumeister's group has repeatedly found that an effort of will or self-control is tiring; if you have had to force yourself to do something, you are less willing or less able to exert self-control when the next challenge comes

around. The phenomenon has been named *ego depletion*. In a typical demonstration, participants who are instructed to stifle their emotional reaction to an emotionally charged film will later perform poorly on a test of physical stamina—how long they can maintain a strong grip on a dynamometer in spite of increasing discomfort. The emotional effort in the first phase of the experiment reduces the ability to withstand the pain of sustained muscle contraction, and ego-depleted people therefore succumb more quickly to the urge to quit. In another experiment, people are first depleted by a task in which they eat virtuous foods such as radishes and celery while resisting the temptation to indulge in chocolate and rich cookies. Later, these people will give up earlier than normal when faced with a difficult cognitive task.

The list of situations and tasks that are now known to deplete self-control is long and varied. All involve conflict and the need to suppress a natural tendency. They include:

- avoiding the thought of white bears
- inhibiting the emotional response to a stirring film
- making a series of choices that involve conflict trying to impress others
- responding kindly to a partner's bad behavior
- interacting with a person of a different race (for prejudiced individuals)

The list of indications of depletion is also highly diverse:

- deviating from one's diet
- overspending on impulsive purchases
- reacting aggressively to provocation
- persisting less time in a handgrip task
- performing poorly in cognitive tasks and logical decision making

The evidence is persuasive: activities that impose high demands on System 2 require self-control, and the exertion of self-control is depleting and unpleasant. Unlike cognitive load, ego depletion is at least in part a loss of motivation. After exerting self-control in one task, you do not feel like making an effort in another, although you could do it if you really had to. In several experiments, people were able to resist the effects of ego depletion when given a strong incentive to do so. In contrast, increasing effort is not an option when you must keep six digits in short-term memory while per-

forming a task. Ego depletion is not the same mental state as cognitive busyness.

The most surprising discovery made by Baumeister's group shows, as he puts it, that the idea of mental energy is more than a mere metaphor. The nervous system consumes more glucose than most other parts of the body, and effortful mental activity appears to be especially expensive in the currency of glucose. When you are actively involved in difficult cognitive reasoning or engaged in a task that requires self-control, your blood glucose level drops. The effect is analogous to a runner who draws down glucose stored in her muscles during a sprint. The bold implication of this idea is that the effects of ego depletion could be undone by ingesting glucose, and Baumeister and his colleagues have confirmed this hypothesis in several experiments.

Volunteers in one of their studies watched a short silent film of a woman being interviewed and were asked to interpret her body language. While they were performing the task, a series of words crossed the screen in slow succession. The participants were specifically instructed to ignore the words, and if they found their attention drawn away they had to refocus their concentration on the woman's behavior. This act of self-control was known to cause ego depletion. All the volunteers drank some lemonade before participating in a second task. The lemonade was sweetened with glucose for half of them and with Splenda for the others. Then all participants were given a task in which they needed to overcome an intuitive response to get the correct answer. Intuitive errors are normally much more frequent among ego-depleted people, and the drinkers of Splenda showed the expected depletion effect. On the other hand, the glucose drinkers were not depleted. Restoring the level of available sugar in the brain had prevented the deterioration of performance. It will take some time and much further research to establish whether the tasks that cause glucose-depletion also cause the momentary arousal that is reflected in increases of pupil size and heart rate.

A disturbing demonstration of depletion effects in judgment was recently reported in the *Proceedings of the National Academy of Sciences*. The unwitting participants in the study were eight parole judges in Israel. They spend entire days reviewing applications for parole. The cases are presented in random order, and the judges spend little time on each one, an average of 6 minutes. (The default decision is denial of parole; only 35% of requests are approved. The exact time of each decision is recorded, and the times of the judges' three food breaks—morning break, lunch, and afternoon break—during the day are recorded as well.) The authors of the study plotted the

proportion of approved requests against the time since the last food break. The proportion spikes after each meal, when about 65% of requests are granted. During the two hours or so until the judges' next feeding, the approval rate drops steadily, to about zero just before the meal. As you might expect, this is an unwelcome result and the authors carefully checked many alternative explanations. The best possible account of the data provides bad news: tired and hungry judges tend to fall back on the easier default position of denying requests for parole. Both fatigue and hunger probably play a role.

THE LAZY SYSTEM 2

One of the main functions of System 2 is to monitor and control thoughts and actions "suggested" by System 1, allowing some to be expressed directly in behavior and suppressing or modifying others.

For an example, here is a simple puzzle. Do not try to solve it but listen to your intuition:

A bat and ball cost \$1.10.
The bat costs one dollar more than the ball.
How much does the ball cost?

A number came to your mind. The number, of course, is 10¢. The distinctive mark of this easy puzzle is that it evokes an answer that is intuitive, appealing, and wrong. Do the math, and you will see. If the ball costs 10¢, then the total cost will be \$1.20 (10¢ for the ball and \$1.10 for the bat), not \$1.10. The correct answer is 5¢. It is safe to assume that the intuitive answer also came to the mind of those who ended up with the correct number—they somehow managed to resist the intuition.

Shane Frederick and I worked together on a theory of judgment based on two systems, and he used the bat-and-ball puzzle to study a central question: How closely does System 2 monitor the suggestions of System 1? His reasoning was that we know a significant fact about anyone who says that the ball costs 10¢: that person did not actively check whether the answer was correct, and her System 2 endorsed an intuitive answer that it could have rejected with a small investment of effort. Furthermore, we also know that the people who give the intuitive answer have missed an obvious social cue; they should have wondered why anyone would include in a questionnaire a puzzle with such an obvious answer. A failure to check is remarkable

because the cost of checking is so low: a few seconds of mental work (the problem is moderately difficult), with slightly tensed muscles and dilated pupils, could avoid an embarrassing mistake. People who say 10¢ appear to be ardent followers of the law of least effort. People who avoid that answer appear to have more active minds.

Many thousands of university students have answered the bat-and-ball puzzle, and the results are shocking. More than 50% of students at Harvard, MIT, and Princeton gave the intuitive—incorrect—answer. At less selective universities, the rate of demonstrable failure to check was in excess of 80%. The bat-and-ball problem is our first encounter with an observation that will be a recurrent theme of this book: many people are overconfident, prone to place too much faith in their intuitions. They apparently find cognitive effort at least mildly unpleasant and avoid it as much as possible.

Now I will show you a logical argument—two premises and a conclusion. Try to determine, as quickly as you can, if the argument is logically valid. Does the conclusion follow from the premises?

All roses are flowers.
Some flowers fade quickly.
Therefore some roses fade quickly.

A large majority of college students endorse this syllogism as valid. In fact the argument is flawed, because it is possible that there are no roses among the flowers that fade quickly. Just as in the bat-and-ball problem, a plausible answer comes to mind immediately. Overriding it requires hard work—the insistent idea that "it's true, it's true!" makes it difficult to check the logic, and most people do not take the trouble to think through the problem.

This experiment has discouraging implications for reasoning in everyday life. It suggests that when people believe a conclusion is true, they are also very likely to believe arguments that appear to support it, even when these arguments are unsound. If System 1 is involved, the conclusion comes first and the arguments follow.

Next, consider the following question and answer it quickly before reading on:

How many murders occur in the state of Michigan in one year?

The question, which was also devised by Shane Frederick, is again a challenge to System 2. The "trick" is whether the respondent will remember that

Detroit, a high-crime city, is in Michigan. College students in the United States know this fact and will correctly identify Detroit as the largest city in Michigan. But knowledge of a fact is not all-or-none. Facts that we know do not always come to mind when we need them. People who remember that Detroit is in Michigan give higher estimates of the murder rate in the state than people who do not, but a majority of Frederick's respondents did not think of the city when questioned about the state. Indeed, the average guess by people who were asked about Michigan is *lower* than the guesses of a similar group who were asked about the murder rate in Detroit.

Blame for a failure to think of Detroit can be laid on both System 1 and System 2. Whether the city comes to mind when the state is mentioned depends in part on the automatic function of memory. People differ in this respect. The representation of the state of Michigan is very detailed in some people's minds: residents of the state are more likely to retrieve many facts about it than people who live elsewhere; geography buffs will retrieve more than others who specialize in baseball statistics; more intelligent individuals are more likely than others to have rich representations of most things. Intelligence is not only the ability to reason; it is also the ability to find relevant material in memory and to deploy attention when needed. Memory function is an attribute of System 1. However, everyone has the option of slowing down to conduct an active search of memory for all possibly relevant facts—just as they could slow down to check the intuitive answer in the bat-and-ball problem. The extent of deliberate checking and search is a characteristic of System 2, which varies among individuals.

The bat-and-ball problem, the flowers syllogism, and the Michigan/Detroit problem have something in common. Failing these minitest appears to be, at least to some extent, a matter of insufficient motivation, not trying hard enough. Anyone who can be admitted to a good university is certainly able to reason through the first two questions and to reflect about Michigan long enough to remember the major city in that state and its crime problem. These students can solve much more difficult problems when they are not tempted to accept a superficially plausible answer that comes readily to mind. The ease with which they are satisfied enough to stop thinking is rather troubling. "Lazy" is a harsh judgment about the self-monitoring of these young people and their System 2, but it does not seem to be unfair. Those who avoid the sin of intellectual sloth could be called "engaged." They are more alert, more intellectually active, less willing to be satisfied with superficially attractive answers, more skeptical about their intuitions. The psychologist Keith Stanovich would call them more rational.

INTELLIGENCE, CONTROL, RATIONALITY

Researchers have applied diverse methods to examine the connection between thinking and self-control. Some have addressed it by asking the correlation question: If people were ranked by their self-control and by their cognitive aptitude, would individuals have similar positions in the two rankings?

In one of the most famous experiments in the history of psychology, Walter Mischel and his students exposed four-year-old children to a cruel dilemma. They were given a choice between a small reward (one Oreo), which they could have at any time, or a larger reward (two cookies) for which they had to wait 15 minutes under difficult conditions. They were to remain alone in a room, facing a desk with two objects: a single cookie and a bell that the child could ring at any time to call in the experimenter and receive the one cookie. As the experiment was described: "There were no toys, books, pictures, or other potentially distracting items in the room. The experimenter left the room and did not return until 15 min had passed or the child had rung the bell, eaten the rewards, stood up, or shown any signs of distress."

The children were watched through a one-way mirror, and the film that shows their behavior during the waiting time always has the audience roaring in laughter. About half the children managed the feat of waiting for 15 minutes, mainly by keeping their attention away from the tempting reward. Ten or fifteen years later, a large gap had opened between those who had resisted temptation and those who had not. The resisters had higher measures of executive control in cognitive tasks, and especially the ability to reallocate their attention effectively. As young adults, they were less likely to take drugs. A significant difference in intellectual aptitude emerged: the children who had shown more self-control as four-year-olds had substantially higher scores on tests of intelligence.

A team of researchers at the University of Oregon explored the link between cognitive control and intelligence in several ways, including an attempt to raise intelligence by improving the control of attention. During five 40-minute sessions, they exposed children aged four to six to various computer games especially designed to demand attention and control. In one of the exercises, the children used a joystick to track a cartoon cat and move it to a grassy area while avoiding a muddy area. The grassy areas gradually shrank and the muddy area expanded, requiring progressively more precise control. The testers found that training attention not only improved

executive control; scores on nonverbal tests of intelligence also improved and the improvement was maintained for several months. Other research by the same group identified specific genes that are involved in the control of attention, showed that parenting techniques also affected this ability, and demonstrated a close connection between the children's ability to control their attention and their ability to control their emotions.

Shane Frederick constructed a Cognitive Reflection Test, which consists of the bat-and-ball problem and two other questions, chosen because they also invite an intuitive answer that is both compelling and wrong (the questions are shown in chapter 5). He went on to study the characteristics of students who score very low on this test—the supervisory function of System 2 is weak in these people—and found that they are prone to answer questions with the first idea that comes to mind and unwilling to invest the effort needed to check their intuitions. Individuals who uncritically follow their intuitions about puzzles are also prone to accept other suggestions from System 1. In particular, they are impulsive, impatient, and keen to receive immediate gratification. For example, 63% of the intuitive respondents say they would prefer to get \$3,400 this month rather than \$3,800 next month. Only 37% of those who solve all three puzzles correctly have the same shortsighted preference for receiving a smaller amount immediately. When asked how much they will pay to get overnight delivery of a book they have ordered, the low scorers on the Cognitive Reflection Test are willing to pay twice as much as the high scorers. Frederick's findings suggest that the characters of our psychodrama have different "personalities." System 1 is impulsive and intuitive; System 2 is capable of reasoning, and it is cautious, but at least for some people it is also lazy. We recognize related differences among individuals: some people are more like their System 2; others are closer to their System 1. This simple test has emerged as one of the better predictors of lazy thinking.

Keith Stanovich and his longtime collaborator Richard West originally introduced the terms System 1 and System 2 (they now prefer to speak of Type 1 and Type 2 processes). Stanovich and his colleagues have spent decades studying differences among individuals in the kinds of problems with which this book is concerned. They have asked one basic question in many different ways: What makes some people more susceptible than others to biases of judgment? Stanovich published his conclusions in a book titled *Rationality and the Reflective Mind*, which offers a bold and distinctive approach to the topic of this chapter. He draws a sharp distinction between two parts of System 2—indeed, the distinction is so sharp that he calls them

separate "minds." One of these minds (he calls it algorithmic) deals with slow thinking and demanding computation. Some people are better than others in these tasks of brain power—they are the individuals who excel in intelligence tests and are able to switch from one task to another quickly and efficiently. However, Stanovich argues that high intelligence does not make people immune to biases. Another ability is involved, which he labels rationality. Stanovich's concept of a rational person is similar to what I earlier labeled "engaged." The core of his argument is that *rationality* should be distinguished from *intelligence*. In his view, superficial or "lazy" thinking is a flaw in the reflective mind, a failure of rationality. This is an attractive and thought-provoking idea. In support of it, Stanovich and his colleagues have found that the bat-and-ball question and others like it are somewhat better indicators of our susceptibility to cognitive errors than are conventional measures of intelligence, such as IQ tests. Time will tell whether the distinction between intelligence and rationality can lead to new discoveries.

SPEAKING OF CONTROL

"She did not have to struggle to stay on task for hours. She was in a state of flow."

"His ego was depleted after a long day of meetings. So he just turned to standard operating procedures instead of thinking through the problem."

"He didn't bother to check whether what he said made sense. Does he usually have a lazy System 2 or was he unusually tired?"

"Unfortunately, she tends to say the first thing that comes into her mind. She probably also has trouble delaying gratification. Weak System 2."

One of my favorite memories of the early years of working with Amos is a comedy routine he enjoyed performing. In a perfect impersonation of one of the professors with whom he had studied philosophy as an undergraduate, Amos would growl in Hebrew marked by a thick German accent: "You must never forget the *Primat of the Is*." What exactly his teacher had meant by that phrase never became clear to me (or to Amos, I believe), but Amos's jokes always made a point. He was reminded of the old phrase (and eventually I was too) whenever we encountered the remarkable asymmetry between the ways our mind treats information that is currently available and information we do not have.

An essential design feature of the associative machine is that it represents only activated ideas. Information that is not retrieved (even unconsciously) from memory might as well not exist. System 1 excels at constructing the best possible story that incorporates ideas currently activated, but it does not (cannot) allow for information it does not have.

The measure of success for System 1 is the coherence of the story it manages to create. The amount and quality of the data on which the story is based are largely irrelevant. When information is scarce, which is a common occurrence, System 1 operates as a machine for jumping to conclusions. Consider the following: "Will Mindik be a good leader? She is intelligent and strong..." An answer quickly came to your mind, and it was yes. You picked the best answer based on the very limited information available, but you jumped the gun. What if the next two adjectives were *corrupt* and *cruel*?

Take note of what you did *not* do as you briefly thought of Mindik as a leader. You did not start by asking, "What would I need to know before I formed an opinion about the quality of someone's leadership?" System 1 got to work on its own from the first adjective: intelligent is good, intelligent and strong is very good. This is the best story that can be constructed from two adjectives, and System 1 delivered it with great cognitive ease. The story will be revised if new information comes in (such as Mindik is corrupt), but there is no waiting and no subjective discomfort. And there also remains a bias favoring the first impression.

The combination of a coherence-seeking System 1 with a lazy System 2 implies that System 2 will endorse many intuitive beliefs, which closely reflect the impressions generated by System 1. Of course, System 2 also is capable of a more systematic and careful approach to evidence, and of following a list of boxes that must be checked before making a decision—think of buying a home, when you deliberately seek information that you don't have. However, System 1 is expected to influence even the more careful decisions. Its input never ceases.

Jumping to conclusions on the basis of limited evidence is so important to an understanding of intuitive thinking, and comes up so often in this book, that I will use a cumbersome abbreviation for it: WYSIATI, which stands for what you see is all there is. System 1 is radically insensitive to both the quality and the quantity of the information that gives rise to impressions and intuitions.

Amos, with two of his graduate students at Stanford, reported a study that bears directly on WYSIATI, by observing the reaction of people who are given one-sided evidence and know it. The participants were exposed to legal scenarios such as the following:

On September 3, plaintiff David Thornton, a forty-three-year-old union field representative, was present in Thrifty Drug Store #168, performing a routine union visit. Within ten minutes of his arrival, a store manager confronted him and told him he could no longer speak with the union employees on the floor of the store. Instead, he would have to see them in a back room while they were on break. Such a request is allowed by the union contract with Thrifty Drug but had never before been enforced. When Mr. Thornton objected, he was told that he had the choice of conforming to these requirements, leaving the store, or being arrested. At this point, Mr. Thornton indicated to the manager that he had always been allowed to speak to employees on the floor for as much as ten minutes, as long as no

business was disrupted, and that he would rather be arrested than change the procedure of his routine visit. The manager then called the police and had Mr. Thornton handcuffed in the store for trespassing. After he was booked and put into a holding cell for a brief time, all charges were dropped. Mr. Thornton is suing Thrifty Drug for false arrest.

In addition to this background material, which all participants read, different groups were exposed to presentations by the lawyers for the two parties. Naturally, the lawyer for the union organizer described the arrest as an intimidation attempt, while the lawyer for the store argued that having the talk in the store was disruptive and that the manager was acting properly. Some participants, like a jury, heard both sides. The lawyers added no useful information that you could not infer from the background story.

The participants were fully aware of the setup, and those who heard only one side could easily have generated the argument for the other side. Nevertheless, the presentation of one-sided evidence had a very pronounced effect on judgments. Furthermore, participants who saw one-sided evidence were more confident of their judgments than those who saw both sides. This is just what you would expect if the confidence that people experience is determined by the coherence of the story they manage to construct from available information. It is the consistency of the information that matters for a good story, not its completeness. Indeed, you will often find that knowing little makes it easier to fit everything you know into a coherent pattern.

WYSIATI facilitates the achievement of coherence and of the cognitive ease that causes us to accept a statement as true. It explains why we can think fast, and how we are able to make sense of partial information in a complex world. Much of the time, the coherent story we put together is close enough to reality to support reasonable action. However, I will also invoke WYSIATI to help explain a long and diverse list of biases of judgment and choice, including the following among many others:

- **Overconfidence:** As the WYSIATI rule implies, neither the quantity nor the quality of the evidence counts for much in subjective confidence. The confidence that individuals have in their beliefs depends mostly on the quality of the story they can tell about what they see, even if they see little. We often fail to allow for the possibility that evidence that should be critical to our judgment is missing—what we see is all there is. Furthermore, our associative system tends to

settle on a coherent pattern of activation and suppresses doubt and ambiguity.

- Framing effects: Different ways of presenting the same information often evoke different emotions. The statement that “the odds of survival one month after surgery are 90%” is more reassuring than the equivalent statement that “mortality within one month of surgery is 10%.” Similarly, cold cuts described as “90% fat-free” are more attractive than when they are described as “10% fat.” The equivalence of the alternative formulations is transparent, but an individual normally sees only one formulation, and what she sees is all there is.
- Base-rate neglect: Recall Steve, the meek and tidy soul who is often believed to be a librarian. The personality description is salient and vivid, and although you surely know that there are more male farmers than male librarians, that statistical fact almost certainly did not come to your mind when you first considered the question. What you saw was all there was.

SPEAKING OF JUMPING TO CONCLUSIONS

“She knows nothing about this person’s management skills. All she is going by is the halo effect from a good presentation.”

“Let’s decorrelate errors by obtaining separate judgments on the issue before any discussion. We will get more information from independent assessments.”

“They made that big decision on the basis of a good report from one consultant. WYSIATI—what you see is all there is. They did not seem to realize how little information they had.”

“They didn’t want more information that might spoil their story. WYSIATI.”

THE ILLUSION OF UNDERSTANDING

The trader-philosopher-statistician Nassim Taleb could also be considered a psychologist. In *The Black Swan*, Taleb introduced the notion of a *narrative fallacy* to describe how flawed stories of the past shape our views of the world and our expectations for the future. Narrative fallacies arise inevitably from our continuous attempt to make sense of the world. The explanatory stories that people find compelling are simple; are concrete rather than abstract; assign a larger role to talent, stupidity, and intentions than to luck; and focus on a few striking events that happened rather than on the countless events that failed to happen. Any recent salient event is a candidate to become the kernel of a causal narrative. Taleb suggests that we humans constantly fool ourselves by constructing flimsy accounts of the past and believing they are true.

Good stories provide a simple and coherent account of people's actions and intentions. You are always ready to interpret behavior as a manifestation of general propensities and personality traits—causes that you can readily match to effects. The halo effect discussed earlier contributes to coherence, because it inclines us to match our view of all the qualities of a person to our judgment of one attribute that is particularly significant. If we think a baseball pitcher is handsome and athletic, for example, we are likely to rate him better at throwing the ball, too. Halos can also be negative: if we think a player is ugly, we will probably underrate his athletic ability. The halo effect helps keep explanatory narratives simple and coherent by exaggerating the consistency of evaluations: good people do only good things and bad people

are all bad. The statement "Hitler loved dogs and little children" is shocking no matter how many times you hear it, because any trace of kindness in someone so evil violates the expectations set up by the halo effect. Inconsistencies reduce the ease of our thoughts and the clarity of our feelings.

A compelling narrative fosters an illusion of inevitability. Consider the story of how Google turned into a giant of the technology industry. Two creative graduate students in the computer science department at Stanford University come up with a superior way of searching information on the Internet. They seek and obtain funding to start a company and make a series of decisions that work out well. Within a few years, the company they started is one of the most valuable stocks in America, and the two former graduate students are among the richest people on the planet. On one memorable occasion, they were lucky, which makes the story even more compelling: a year after founding Google, they were willing to sell their company for less than \$1 million, but the buyer said the price was too high. Mentioning the single lucky incident actually makes it easier to underestimate the multitude of ways in which luck affected the outcome.

A detailed history would specify the decisions of Google's founders, but for our purposes it suffices to say that almost every choice they made had a good outcome. A more complete narrative would describe the actions of the firms that Google defeated. The hapless competitors would appear to be blind, slow, and altogether inadequate in dealing with the threat that eventually overwhelmed them.

I intentionally told this tale blandly, but you get the idea: there is a very good story here. Fleshed out in more detail, the story could give you the sense that you understand what made Google succeed; it would also make you feel that you have learned a valuable general lesson about what makes businesses succeed. Unfortunately, there is good reason to believe that your sense of understanding and learning from the Google story is largely illusory. The ultimate test of an explanation is whether it would have made the event predictable in advance. No story of Google's unlikely success will meet that test, because no story can include the myriad of events that would have caused a different outcome. The human mind does not deal well with nonevents. The fact that many of the important events that did occur involve choices further tempts you to exaggerate the role of skill and underestimate the part that luck played in the outcome. Because every critical decision turned out well, the record suggests almost flawless prescience—but bad luck could have disrupted any one of the successful steps. The halo effect adds the final touches, lending an aura of invincibility to the heroes of the story.

Like watching a skilled rafter avoiding one potential calamity after another as he goes down the rapids, the unfolding of the Google story is thrilling because of the constant risk of disaster. However, there is an instructive difference between the two cases. The skilled rafter has gone down rapids hundreds of times. He has learned to read the roiling water in front of him and to anticipate obstacles. He has learned to make the tiny adjustments of posture that keep him upright. There are fewer opportunities for young men to learn how to create a giant company, and fewer chances to avoid hidden rocks—such as a brilliant innovation by a competing firm. Of course there was a great deal of skill in the Google story, but luck played a more important role in the actual event than it does in the telling of it. And the more luck was involved, the less there is to be learned.

At work here is that powerful WYSIATI rule. You cannot help dealing with the limited information you have as if it were all there is to know. You build the best possible story from the information available to you, and if it is a good story, you believe it. Paradoxically, it is easier to construct a coherent story when you know little, when there are fewer pieces to fit into the puzzle. Our comforting conviction that the world makes sense rests on a secure foundation: our almost unlimited ability to ignore our ignorance.

I have heard of too many people who "knew well before it happened that the 2008 financial crisis was inevitable." This sentence contains a highly objectionable word, which should be removed from our vocabulary in discussions of major events. The word is, of course, *knew*. Some people thought well in advance that there would be a crisis, but they did not know it. They now say they knew it because the crisis did in fact happen. This is a misuse of an important concept. In everyday language, we apply the word *know* only when what was known is true and can be shown to be true. We can know something only if it is both true and knowable. But the people who thought there would be a crisis (and there are fewer of them than now remember thinking it) could not conclusively show it at the time. Many intelligent and well-informed people were keenly interested in the future of the economy and did not believe a catastrophe was imminent; I infer from this fact that the crisis was not knowable. What is perverse about the use of *know* in this context is not that some individuals get credit for prescience that they do not deserve. It is that the language implies that the world is more knowable than it is. It helps perpetuate a pernicious illusion.

The core of the illusion is that we believe we understand the past, which implies that the future also should be knowable, but in fact we understand the past less than we believe we do. *Know* is not the only word that fosters

this illusion. In common usage, the words *intuition* and *premonition* also are reserved for past thoughts that turned out to be true. The statement “I had a premonition that the marriage would not last, but I was wrong” sounds odd, as does any sentence about an intuition that turned out to be false. To think clearly about the future, we need to clean up the language that we use in labeling the beliefs we had in the past.

THE SOCIAL COSTS OF HINDSIGHT

The mind that makes up narratives about the past is a sense-making organ. When an unpredicted event occurs, we immediately adjust our view of the world to accommodate the surprise. Imagine yourself before a football game between two teams that have the same record of wins and losses. Now the game is over, and one team trashed the other. In your revised model of the world, the winning team is much stronger than the loser, and your view of the past as well as of the future has been altered by that new perception. Learning from surprises is a reasonable thing to do, but it can have some dangerous consequences.

A general limitation of the human mind is its imperfect ability to reconstruct past states of knowledge, or beliefs that have changed. Once you adopt a new view of the world (or of any part of it), you immediately lose much of your ability to recall what you used to believe before your mind changed.

Many psychologists have studied what happens when people change their minds. Choosing a topic on which minds are not completely made up—say, the death penalty—the experimenter carefully measures people’s attitudes. Next, the participants see or hear a persuasive pro or con message. Then the experimenter measures people’s attitudes again; they usually are closer to the persuasive message they were exposed to. Finally, the participants report the opinion they held beforehand. This task turns out to be surprisingly difficult. Asked to reconstruct their former beliefs, people retrieve their current ones instead—an instance of substitution—and many cannot believe that they ever felt differently.

Your inability to reconstruct past beliefs will inevitably cause you to underestimate the extent to which you were surprised by past events. Baruch Fischhoff first demonstrated this “I-knew-it-all-along” effect, or *hindsight bias*, when he was a student in Jerusalem. Together with Ruth Beyth (another of our students), Fischhoff conducted a survey before President Richard Nixon visited China and Russia in 1972. The respondents assigned probabilities to fifteen possible outcomes of Nixon’s diplomatic

initiatives. Would Mao Zedong agree to meet with Nixon? Might the United States grant diplomatic recognition to China? After decades of enmity, could the United States and the Soviet Union agree on anything significant?

After Nixon’s return from his travels, Fischhoff and Beyth asked the same people to recall the probability that they had originally assigned to each of the fifteen possible outcomes. The results were clear. If an event had actually occurred, people exaggerated the probability that they had assigned to it earlier. If the possible event had not come to pass, the participants erroneously recalled that they had always considered it unlikely. Further experiments showed that people were driven to overstate the accuracy not only of their original predictions but also of those made by others. Similar results have been found for other events that gripped public attention, such as the O. J. Simpson murder trial and the impeachment of President Bill Clinton. The tendency to revise the history of one’s beliefs in light of what actually happened produces a robust cognitive illusion.

Hindsight bias has pernicious effects on the evaluations of decision makers. It leads observers to assess the quality of a decision not by whether the process was sound but by whether its outcome was good or bad. Consider a low-risk surgical intervention in which an unpredictable accident occurred that caused the patient’s death. The jury will be prone to believe, after the fact, that the operation was actually risky and that the doctor who ordered it should have known better. This outcome bias makes it almost impossible to evaluate a decision properly—in terms of the beliefs that were reasonable when the decision was made.

Hindsight is especially unkind to decision makers who act as agents for others—physicians, financial advisers, third-base coaches, CEOs, social workers, diplomats, politicians. We are prone to blame decision makers for good decisions that worked out badly and to give them too little credit for successful moves that appear obvious only after the fact. There is a clear *outcome bias*. When the outcomes are bad, the clients often blame their agents for not seeing the handwriting on the wall—forgetting that it was written in invisible ink that became legible only afterward. Actions that seemed prudent in foresight can look irresponsibly negligent in hindsight. Based on an actual legal case, students in California were asked whether the city of Duluth, Minnesota, should have shouldered the considerable cost of hiring a full-time bridge monitor to protect against the risk that debris might get caught and block the free flow of water. One group was shown only the evidence available at the time of the city’s decision; 24% of these people felt that Duluth should take on the expense of hiring a flood mon-

itor. The second group was informed that debris had blocked the river, causing major flood damage; 56% of these people said the city should have hired the monitor, although they had been explicitly instructed not to let hindsight distort their judgment.

The worse the consequence, the greater the hindsight bias. In the case of a catastrophe, such as 9/11, we are especially ready to believe that the officials who failed to anticipate it were negligent or blind. On July 10, 2001, the Central Intelligence Agency obtained information that al-Qaeda might be planning a major attack against the United States. George Tenet, director of the CIA, brought the information not to President George W. Bush but to National Security Adviser Condoleezza Rice. When the facts later emerged, Ben Bradlee, the legendary executive editor of *The Washington Post*, declared, "It seems to me elementary that if you've got the story that's going to dominate history you might as well go right to the president." But on July 10, no one knew—or could have known—that this tidbit of intelligence would turn out to dominate history.

Because adherence to standard operating procedures is difficult to second-guess, decision makers who expect to have their decisions scrutinized with hindsight are driven to bureaucratic solutions—and to an extreme reluctance to take risks. As malpractice litigation became more common, physicians changed their procedures in multiple ways: ordered more tests, referred more cases to specialists, applied conventional treatments even when they were unlikely to help. These actions protected the physicians more than they benefited the patients, creating the potential for conflicts of interest. Increased accountability is a mixed blessing.

Although hindsight and the outcome bias generally foster risk aversion, they also bring undeserved rewards to irresponsible risk seekers, such as a general or an entrepreneur who took a crazy gamble and won. Leaders who have been lucky are never punished for having taken too much risk. Instead, they are believed to have had the flair and foresight to anticipate success, and the sensible people who doubted them are seen in hindsight as mediocre, timid, and weak. A few lucky gambles can crown a reckless leader with a halo of prescience and boldness.

RECIPES FOR SUCCESS

The sense-making machinery of System 1 makes us see the world as more tidy, simple, predictable, and coherent than it really is. The illusion that one has understood the past feeds the further illusion that one can predict and

control the future. These illusions are comforting. They reduce the anxiety that we would experience if we allowed ourselves to fully acknowledge the uncertainties of existence. We all have a need for the reassuring message that actions have appropriate consequences, and that success will reward wisdom and courage. Many business books are tailor-made to satisfy this need.

Do leaders and management practices influence the outcomes of firms in the market? Of course they do, and the effects have been confirmed by systematic research that objectively assessed the characteristics of CEOs and their decisions, and related them to subsequent outcomes of the firm. In one study, the CEOs were characterized by the strategy of the companies they had led before their current appointment, as well as by management rules and procedures adopted after their appointment. CEOs do influence performance, but the effects are much smaller than a reading of the business press suggests.

Researchers measure the strength of relationships by a correlation coefficient, which varies between 0 and 1. The coefficient was defined earlier (in relation to regression to the mean) by the extent to which two measures are determined by shared factors. A very generous estimate of the correlation between the success of the firm and the quality of its CEO might be as high as .30, indicating 30% overlap. To appreciate the significance of this number, consider the following question:

Suppose you consider many pairs of firms. The two firms in each pair are generally similar, but the CEO of one of them is better than the other. How often will you find that the firm with the stronger CEO is the more successful of the two?

In a well-ordered and predictable world, the correlation would be perfect (1), and the stronger CEO would be found to lead the more successful firm in 100% of the pairs. If the relative success of similar firms was determined entirely by factors that the CEO does not control (call them luck, if you wish), you would find the more successful firm led by the weaker CEO 50% of the time. A correlation of .30 implies that you would find the stronger CEO leading the stronger firm in about 60% of the pairs—an improvement of a mere 10 percentage points over random guessing, hardly grist for the hero worship of CEOs we so often witness.

If you expected this value to be higher—and most of us do—then you should take that as an indication that you are prone to overestimate the

predictability of the world you live in. Make no mistake: improving the odds of success from 1:1 to 3:2 is a very significant advantage, both at the racetrack and in business. From the perspective of most business writers, however, a CEO who has so little control over performance would not be particularly impressive even if her firm did well. It is difficult to imagine people lining up at airport bookstores to buy a book that enthusiastically describes the practices of business leaders who, on average, do somewhat better than chance. Consumers have a hunger for a clear message about the determinants of success and failure in business, and they need stories that offer a sense of understanding, however illusory.

In his penetrating book *The Halo Effect*, Philip Rosenzweig, a business school professor based in Switzerland, shows how the demand for illusory certainty is met in two popular genres of business writing: histories of the rise (usually) and fall (occasionally) of particular individuals and companies, and analyses of differences between successful and less successful firms. He concludes that stories of success and failure consistently exaggerate the impact of leadership style and management practices on firm outcomes, and thus their message is rarely useful.

To appreciate what is going on, imagine that business experts, such as other CEOs, are asked to comment on the reputation of the chief executive of a company. They are keenly aware of whether the company has recently been thriving or failing. As we saw earlier in the case of Google, this knowledge generates a halo. The CEO of a successful company is likely to be called flexible, methodical, and decisive. Imagine that a year has passed and things have gone sour. The same executive is now described as confused, rigid, and authoritarian. Both descriptions sound right at the time: it seems almost absurd to call a successful leader rigid and confused, or a struggling leader flexible and methodical.

Indeed, the halo effect is so powerful that you probably find yourself resisting the idea that the same person and the same behaviors appear methodical when things are going well and rigid when things are going poorly. Because of the halo effect, we get the causal relationship backward: we are prone to believe that the firm fails because its CEO is rigid, when the truth is that the CEO appears to be rigid because the firm is failing. This is how illusions of understanding are born.

The halo effect and outcome bias combine to explain the extraordinary appeal of books that seek to draw operational morals from systematic examination of successful businesses. One of the best-known examples of this genre is Jim Collins and Jerry I. Porras's *Built to Last*. The book contains a

thorough analysis of eighteen pairs of competing companies, in which one was more successful than the other. The data for these comparisons are ratings of various aspects of corporate culture, strategy, and management practices. "We believe every CEO, manager, and entrepreneur in the world should read this book," the authors proclaim. "You can build a visionary company."

The basic message of *Built to Last* and other similar books is that good managerial practices can be identified and that good practices will be rewarded by good results. Both messages are overstated. The comparison of firms that have been more or less successful is to a significant extent a comparison between firms that have been more or less lucky. Knowing the importance of luck, you should be particularly suspicious when highly consistent patterns emerge from the comparison of successful and less successful firms. In the presence of randomness, regular patterns can only be mirages.

Because luck plays a large role, the quality of leadership and management practices cannot be inferred reliably from observations of success. And even if you had perfect foreknowledge that a CEO has brilliant vision and extraordinary competence, you still would be unable to predict how the company will perform with much better accuracy than the flip of a coin. On average, the gap in corporate profitability and stock returns between the outstanding firms and the less successful firms studied in *Built to Last* shrank to almost nothing in the period following the study. The average profitability of the companies identified in the famous *In Search of Excellence* dropped sharply as well within a short time. A study of *Fortune's* "Most Admired Companies" finds that over a twenty-year period, the firms with the worst ratings went on to earn much higher stock returns than the most admired firms.

You are probably tempted to think of causal explanations for these observations: perhaps the successful firms became complacent, the less successful firms tried harder. But this is the wrong way to think about what happened. The average gap must shrink, because the original gap was due in good part to luck, which contributed both to the success of the top firms and to the lagging performance of the rest. We have already encountered this statistical fact of life: regression to the mean.

Stories of how businesses rise and fall strike a chord with readers by offering what the human mind needs: a simple message of triumph and failure that identifies clear causes and ignores the determinative power of luck and the inevitability of regression. These stories induce and maintain an illusion

of understanding, imparting lessons of little enduring value to readers who are all too eager to believe them.

SPEAKING OF HINDSIGHT

"The mistake appears obvious, but it is just hindsight. You could not have known in advance."

"He's learning too much from this success story, which is too tidy. He has fallen for a narrative fallacy."

"She has no evidence for saying that the firm is badly managed. All she knows is that its stock has gone down. This is an outcome bias, part hindsight and part halo effect."

"Let's not fall for the outcome bias. This was a stupid decision even though it worked out well."

THE ILLUSION OF VALIDITY

System 1 is designed to jump to conclusions from little evidence—and it is not designed to know the size of its jumps. Because of WYSIATI, only the evidence at hand counts. Because of confidence by coherence, the subjective confidence we have in our opinions reflects the coherence of the story that System 1 and System 2 have constructed. The amount of evidence and its quality do not count for much, because poor evidence can make a very good story. For some of our most important beliefs we have no evidence at all, except that people we love and trust hold these beliefs. Considering how little we know, the confidence we have in our beliefs is preposterous—and it is also essential.

THE ILLUSION OF VALIDITY

Many decades ago I spent what seemed like a great deal of time under a scorching sun, watching groups of sweaty soldiers as they solved a problem. I was doing my national service in the Israeli Army at the time. I had completed an undergraduate degree in psychology, and after a year as an infantry officer was assigned to the army's Psychology Branch, where one of my occasional duties was to help evaluate candidates for officer training. We used methods that had been developed by the British Army in World War II.

One test, called the "leaderless group challenge," was conducted on an obstacle field. Eight candidates, strangers to each other, with all insignia of rank removed and only numbered tags to identify them, were instructed to

ilar results across a variety of other forecast outcomes, including violations of parole, success in pilot training, and criminal recidivism.

Not surprisingly, Meehl's book provoked shock and disbelief among clinical psychologists, and the controversy it started has engendered a stream of research that is still flowing today, more than fifty years after its publication. The number of studies reporting comparisons of clinical and statistical predictions has increased to roughly two hundred, but the score in the contest between algorithms and humans has not changed. About 60% of the studies have shown significantly better accuracy for the algorithms. The other comparisons scored a draw in accuracy, but a tie is tantamount to a win for the statistical rules, which are normally much less expensive to use than expert judgment. No exception has been convincingly documented.

The range of predicted outcomes has expanded to cover medical variables such as the longevity of cancer patients, the length of hospital stays, the diagnosis of cardiac disease, and the susceptibility of babies to sudden infant death syndrome; economic measures such as the prospects of success for new businesses, the evaluation of credit risks by banks, and the future career satisfaction of workers; questions of interest to government agencies, including assessments of the suitability of foster parents, the odds of recidivism among juvenile offenders, and the likelihood of other forms of violent behavior; and miscellaneous outcomes such as the evaluation of scientific presentations, the winners of football games, and the future prices of Bordeaux wine. Each of these domains entails a significant degree of uncertainty and unpredictability. We describe them as "low-validity environments." In every case, the accuracy of experts was matched or exceeded by a simple algorithm.

As Meehl pointed out with justified pride thirty years after the publication of his book, "There is no controversy in social science which shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one."

The Princeton economist and wine lover Orley Ashenfelter has offered a compelling demonstration of the power of simple statistics to outdo world-renowned experts. Ashenfelter wanted to predict the future value of fine Bordeaux wines from information available in the year they are made. The question is important because fine wines take years to reach their peak quality, and the prices of mature wines from the same vineyard vary dramatically across different vintages; bottles filled only twelve months apart

INTUITIONS VS. FORMULAS

Paul Meehl was a strange and wonderful character, and one of the most versatile psychologists of the twentieth century. Among the departments in which he had faculty appointments at the University of Minnesota were psychology, law, psychiatry, neurology, and philosophy. He also wrote on religion, political science, and learning in rats. A statistically sophisticated researcher and a fierce critic of empty claims in clinical psychology, Meehl was also a practicing psychoanalyst. He wrote thoughtful essays on the philosophical foundations of psychological research that I almost memorized while I was a graduate student. I never met Meehl, but he was one of my heroes from the time I read his *Clinical vs. Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*.

In the slim volume that he later called "my disturbing little book," Meehl reviewed the results of 20 studies that had analyzed whether *clinical predictions* based on the subjective impressions of trained professionals were more accurate than *statistical predictions* made by combining a few scores or ratings according to a rule. In a typical study, trained counselors predicted the grades of freshmen at the end of the school year. The counselors interviewed each student for forty-five minutes. They also had access to high school grades, several aptitude tests, and a four-page personal statement. The statistical algorithm used only a fraction of this information: high school grades and one aptitude test. Nevertheless, the formula was more accurate than 11 of the 14 counselors. Meehl reported generally sim-

can differ in value by a factor of 10 or more. An ability to forecast future prices is of substantial value, because investors buy wine, like art, in the anticipation that its value will appreciate.

It is generally agreed that the effect of vintage can be due only to variations in the weather during the grape-growing season. The best wines are produced when the summer is warm and dry, which makes the Bordeaux wine industry a likely beneficiary of global warming. The industry is also helped by wet springs, which increase quantity without much effect on quality. Ashenfelter converted that conventional knowledge into a statistical formula that predicts the price of a wine—for a particular property and at a particular age—by three features of the weather: the average temperature over the summer growing season, the amount of rain at harvest time, and the total rainfall during the previous winter. His formula provides accurate price forecasts years and even decades into the future. Indeed, his formula forecasts future prices much more accurately than the current prices of young wines do. This new example of a “Meehl pattern” challenges the abilities of the experts whose opinions help shape the early price. It also challenges economic theory, according to which prices should reflect all the available information, including the weather. Ashenfelter’s formula is extremely accurate—the correlation between his predictions and actual prices is above .90.

Why are experts inferior to algorithms? One reason, which Meehl suspected, is that experts try to be clever, think outside the box, and consider complex combinations of features in making their predictions. Complexity may work in the odd case, but more often than not it reduces validity. Simple combinations of features are better. Several studies have shown that human decision makers are inferior to a prediction formula even when they are given the score suggested by the formula! They feel that they can overrule the formula because they have additional information about the case, but they are wrong more often than not. According to Meehl, there are few circumstances under which it is a good idea to substitute judgment for a formula. In a famous thought experiment, he described a formula that predicts whether a particular person will go to the movies tonight and noted that it is proper to disregard the formula if information is received that the individual broke a leg today. The name “broken-leg rule” has stuck. The point, of course, is that broken legs are very rare—as well as decisive.

Another reason for the inferiority of expert judgment is that humans are incorrigibly inconsistent in making summary judgments of complex information. When asked to evaluate the same information twice, they fre-

quently give different answers. The extent of the inconsistency is often a matter of real concern. Experienced radiologists who evaluate chest X-rays as “normal” or “abnormal” contradict themselves 20% of the time when they see the same picture on separate occasions. A study of 101 independent auditors who were asked to evaluate the reliability of internal corporate audits revealed a similar degree of inconsistency. A review of 41 separate studies of the reliability of judgments made by auditors, pathologists, psychologists, organizational managers, and other professionals suggests that this level of inconsistency is typical, even when a case is reevaluated within a few minutes. Unreliable judgments cannot be valid predictors of anything.

The widespread inconsistency is probably due to the extreme context dependency of System 1. We know from studies of priming that unnoticed stimuli in our environment have a substantial influence on our thoughts and actions. These influences fluctuate from moment to moment. The brief pleasure of a cool breeze on a hot day may make you slightly more positive and optimistic about whatever you are evaluating at the time. The prospects of a convict being granted parole may change significantly during the time that elapses between successive food breaks in the parole judges’ schedule. Because you have little direct knowledge of what goes on in your mind, you will never know that you might have made a different judgment or reached a different decision under very slightly different circumstances. Formulas do not suffer from such problems. Given the same input, they always return the same answer. When predictability is poor—which it is in most of the studies reviewed by Meehl and his followers—inconsistency is destructive of any predictive validity.

The research suggests a surprising conclusion: to maximize predictive accuracy, final decisions should be left to formulas, especially in low-validity environments. In admission decisions for medical schools, for example, the final determination is often made by the faculty members who interview the candidate. The evidence is fragmentary, but there are solid grounds for a conjecture: conducting an interview is likely to diminish the accuracy of a selection procedure, if the interviewers also make the final admission decisions. Because interviewers are overconfident in their intuitions, they will assign too much weight to their personal impressions and too little weight to other sources of information, lowering validity. Similarly, the experts who evaluate the quality of immature wine to predict its future have a source of information that almost certainly makes things worse rather than better: they can taste the wine. In addition, of course, even if they have a

good understanding of the effects of the weather on wine quality, they will not be able to maintain the consistency of a formula.

The most important development in the field since Meehl's original work is Robyn Dawes's famous article "The Robust Beauty of Improper Linear Models in Decision Making." The dominant statistical practice in the social sciences is to assign weights to the different predictors by following an algorithm, called multiple regression, that is now built into conventional software. The logic of multiple regression is unassailable: it finds the optimal formula for putting together a weighted combination of the predictors. However, Dawes observed that the complex statistical algorithm adds little or no value. One can do just as well by selecting a set of scores that have some validity for predicting the outcome and adjusting the values to make them comparable (by using standard scores or ranks). A formula that combines these predictors with equal weights is likely to be just as accurate in predicting new cases as the multiple-regression formula that was optimal in the original sample. More recent research went further: formulas that assign equal weights to all the predictors are often superior, because they are not affected by accidents of sampling.

The surprising success of equal-weighting schemes has an important practical implication: it is possible to develop useful algorithms without any prior statistical research. Simple equally weighted formulas based on existing statistics or on common sense are often very good predictors of significant outcomes. In a memorable example, Dawes showed that marital stability is well predicted by a formula:

frequency of lovemaking minus frequency of quarrels

You don't want your result to be a negative number.

The important conclusion from this research is that an algorithm that is constructed on the back of an envelope is often good enough to compete with an optimally weighted formula, and certainly good enough to outdo expert judgment. This logic can be applied in many domains, ranging from the selection of stocks by portfolio managers to the choices of medical treatments by doctors or patients.

A classic application of this approach is a simple algorithm that has saved the lives of hundreds of thousands of infants. Obstetricians had always known that an infant who is not breathing normally within a few

minutes of birth is at high risk of brain damage or death. Until the anesthesiologist Virginia Apgar intervened in 1953, physicians and midwives used their clinical judgment to determine whether a baby was in distress. Different practitioners focused on different cues. Some watched for breathing problems while others monitored how soon the baby cried. Without a standardized procedure, danger signs were often missed, and many newborn infants died.

One day over breakfast, a medical resident asked how Dr. Apgar would make a systematic assessment of a newborn. "That's easy," she replied. "You would do it like this." Apgar jotted down five variables (heart rate, respiration, reflex, muscle tone, and color) and three scores (0, 1, or 2, depending on the robustness of each sign). Realizing that she might have made a breakthrough that any delivery room could implement, Apgar began rating infants by this rule one minute after they were born. A baby with a total score of 8 or above was likely to be pink, squirming, crying, grimacing, with a pulse of 100 or more—in good shape. A baby with a score of 4 or below was probably bluish, flaccid, passive, with a slow or weak pulse—in need of immediate intervention. Applying Apgar's score, the staff in delivery rooms finally had consistent standards for determining which babies were in trouble, and the formula is credited for an important contribution to reducing infant mortality. The Apgar test is still used every day in every delivery room. Atul Gawande's recent *A Checklist Manifesto* provides many other examples of the virtues of checklists and simple rules.

THE HOSTILITY TO ALGORITHMS

From the very outset, clinical psychologists responded to Meehl's ideas with hostility and disbelief. Clearly, they were in the grip of an illusion of skill in terms of their ability to make long-term predictions. On reflection, it is easy to see how the illusion came about and easy to sympathize with the clinicians' rejection of Meehl's research.

The statistical evidence of clinical inferiority contradicts clinicians' everyday experience of the quality of their judgments. Psychologists who work with patients have many hunches during each therapy session, anticipating how the patient will respond to an intervention, guessing what will happen next. Many of these hunches are confirmed, illustrating the reality of clinical skill.

The problem is that the correct judgments involve short-term predictions in the context of the therapeutic interview, a skill in which therapists may have years of practice. The tasks at which they fail typically require long-term predictions about the patient's future. These are much more difficult, even the best formulas do only modestly well, and they are also tasks that the clinicians have never had the opportunity to learn properly—they would have to wait years for feedback, instead of receiving the instantaneous feedback of the clinical session. However, the line between what clinicians can do well and what they cannot do at all well is not obvious, and certainly not obvious to them. They know they are skilled, but they don't necessarily know the boundaries of their skill. Not surprisingly, then, the idea that a mechanical combination of a few variables could outperform the subtle complexity of human judgment strikes experienced clinicians as obviously wrong.

The debate about the virtues of clinical and statistical prediction has always had a moral dimension. The statistical method, Meehl wrote, was criticized by experienced clinicians as "mechanical, atomistic, additive, cut and dried, artificial, unreal, arbitrary, incomplete, dead, pedantic, fractionated, trivial, forced, static, superficial, rigid, sterile, academic, pseudoscientific and blind." The clinical method, on the other hand, was lauded by its proponents as "dynamic, global, meaningful, holistic, subtle, sympathetic, configural, patterned, organized, rich, deep, genuine, sensitive, sophisticated, real, living, concrete, natural, true to life, and understanding."

This is an attitude we can all recognize. When a human competes with a machine, whether it is John Henry a-hammerin' on the mountain or the chess genius Garry Kasparov facing off against the computer Deep Blue, our sympathies lie with our fellow human. The aversion to algorithms making decisions that affect humans is rooted in the strong preference that many people have for the natural over the synthetic or artificial. Asked whether they would rather eat an organic or a commercially grown apple, most people prefer the "all natural" one. Even after being informed that the two apples taste the same, have identical nutritional value, and are equally healthful, a majority still prefer the organic fruit. Even the producers of beer have found that they can increase sales by putting "All Natural" or "No Preservatives" on the label.

The deep resistance to the demystification of expertise is illustrated by the reaction of the European wine community to Ashenfelter's formula for predicting the price of Bordeaux wines. Ashenfelter's formula answered a prayer: one might thus have expected that wine lovers everywhere would be grateful to him for demonstrably improving their ability to identify the

wines that later would be good. Not so. The response in French wine circles, wrote *The New York Times*, ranged "somewhere between violent and hysterical." Ashenfelter reports that one oenophile called his findings "ludicrous and absurd." Another scoffed, "It is like judging movies without actually seeing them."

The prejudice against algorithms is magnified when the decisions are consequential. Meehl remarked, "I do not quite know how to alleviate the horror some clinicians seem to experience when they envisage a treatable case being denied treatment because a 'blind, mechanical' equation misclassifies him." In contrast, Meehl and other proponents of algorithms have argued strongly that it is unethical to rely on intuitive judgments for important decisions if an algorithm is available that will make fewer mistakes. Their rational argument is compelling, but it runs against a stubborn psychological reality: for most people, the cause of a mistake matters. The story of a child dying because an algorithm made a mistake is more poignant than the story of the same tragedy occurring as a result of human error, and the difference in emotional intensity is readily translated into a moral preference.

Fortunately, the hostility to algorithms will probably soften as their role in everyday life continues to expand. Looking for books or music we might enjoy, we appreciate recommendations generated by software. We take it for granted that decisions about credit limits are made without the direct intervention of any human judgment. We are increasingly exposed to guidelines that have the form of simple algorithms, such as the ratio of good and bad cholesterol levels we should strive to attain. The public is now well aware that formulas may do better than humans in some critical decisions in the world of sports: how much a professional team should pay for particular rookie players, or when to punt on fourth down. The expanding list of tasks that are assigned to algorithms should eventually reduce the discomfort that most people feel when they first encounter the pattern of results that Meehl described in his disturbing little book.

LEARNING FROM MEEHL

In 1955, as a twenty-one-year-old lieutenant in the Israeli Defense Forces, I was assigned to set up an interview system for the entire army. If you wonder why such a responsibility would be forced upon someone so young, bear in mind that the state of Israel itself was only seven years old at the time; all its institutions were under construction, and someone had to build them. Odd

as it sounds today, my bachelor's degree in psychology probably qualified me as the best-trained psychologist in the army. My direct supervisor, a brilliant researcher, had a degree in chemistry.

An interview routine was already in place when I was given my mission. Every soldier drafted into the army completed a battery of psychometric tests, and each man considered for combat duty was interviewed for an assessment of personality. The goal was to assign the recruit a score of general fitness for combat and to find the best match of his personality among various branches: infantry, artillery, armor, and so on. The interviewers were themselves young draftees, selected for this assignment by virtue of their high intelligence and interest in dealing with people. Most were women, who were at the time exempt from combat duty. Trained for a few weeks in how to conduct a fifteen- to twenty-minute interview, they were encouraged to cover a range of topics and to form a general impression of how well the recruit would do in the army.

Unfortunately, follow-up evaluations had already indicated that this interview procedure was almost useless for predicting the future success of recruits. I was instructed to design an interview that would be more useful but would not take more time. I was also told to try out the new interview and to evaluate its accuracy. From the perspective of a serious professional, I was no more qualified for the task than I was to build a bridge across the Amazon.

Fortunately, I had read Paul Meehl's "little book," which had appeared just a year earlier. I was convinced by his argument that simple, statistical rules are superior to intuitive "clinical" judgments. I concluded that the then current interview had failed at least in part because it allowed the interviewers to do what they found most interesting, which was to learn about the dynamics of the interviewee's mental life. Instead, we should use the limited time at our disposal to obtain as much specific information as possible about the interviewee's life in his normal environment. Another lesson I learned from Meehl was that we should abandon the procedure in which the interviewers' global evaluations of the recruit determined the final decision. Meehl's book suggested that such evaluations should not be trusted and that statistical summaries of separately evaluated attributes would achieve higher validity.

I decided on a procedure in which the interviewers would evaluate several relevant personality traits and score each separately. The final score of fitness for combat duty would be computed according to a standard formula, with no further input from the interviewers. I made up a list of six charac-

teristics that appeared relevant to performance in a combat unit, including "responsibility," "sociability," and "masculine pride." I then composed, for each trait, a series of factual questions about the individual's life before his enlistment, including the number of different jobs he had held, how regular and punctual he had been in his work or studies, the frequency of his interactions with friends, and his interest and participation in sports, among others. The idea was to evaluate as objectively as possible how well the recruit had done on each dimension.

By focusing on standardized, factual questions, I hoped to combat the halo effect, where favorable first impressions influence later judgments. As a further precaution against halos, I instructed the interviewers to go through the six traits in a fixed sequence, rating each trait on a five-point scale before going on to the next. And that was that. I informed the interviewers that they need not concern themselves with the recruit's future adjustment to the military. Their only task was to elicit relevant facts about his past and to use that information to score each personality dimension. "Your function is to provide reliable measurements," I told them. "Leave the predictive validity to me," by which I meant the formula that I was going to devise to combine their specific ratings.

The interviewers came close to mutiny. These bright young people were displeased to be ordered, by someone hardly older than themselves, to switch off their intuition and focus entirely on boring factual questions. One of them complained, "You are turning us into robots!" So I compromised. "Carry out the interview exactly as instructed," I told them, "and when you are done, have your wish: close your eyes, try to imagine the recruit as a soldier, and assign him a score on a scale of 1 to 5."

Several hundred interviews were conducted by this new method, and a few months later we collected evaluations of the soldiers' performance from the commanding officers of the units to which they had been assigned. The results made us happy. As Meehl's book had suggested, the new interview procedure was a substantial improvement over the old one. The sum of our six ratings predicted soldiers' performance much more accurately than the global evaluations of the previous interviewing method, although far from perfectly. We had progressed from "completely useless" to "moderately useful."

The big surprise to me was that the intuitive judgment that the interviewers summoned up in the "close your eyes" exercise also did very well, indeed just as well as the sum of the six specific ratings. I learned from this finding a lesson that I have never forgotten: intuition adds value even in the

justly derided selection interview, but only after a disciplined collection of objective information and disciplined scoring of separate traits. I set a formula that gave the “close your eyes” evaluation the same weight as the sum of the six trait ratings. A more general lesson that I learned from this episode was do not simply trust intuitive judgment—your own or that of others—but do not dismiss it, either.

Some forty-five years later, after I won a Nobel Prize in economics, I was for a short time a minor celebrity in Israel. On one of my visits, someone had the idea of escorting me around my old army base, which still housed the unit that interviewed new recruits. I was introduced to the commanding officer of the Psychological Unit, and she described their current interviewing practices, which had not changed much from the system I had designed; there was, it turned out, a considerable amount of research indicating that the interviews still worked well. As she came to the end of her description of how the interviews are conducted, the officer added, “And then we tell them, ‘Close your eyes.’”

DO IT YOURSELF

The message of this chapter is readily applicable to tasks other than making manpower decisions for an army. Implementing interview procedures in the spirit of Meehl and Dawes requires relatively little effort but substantial discipline. Suppose that you need to hire a sales representative for your firm. If you are serious about hiring the best possible person for the job, this is what you should do. First, select a few traits that are prerequisites for success in this position (technical proficiency, engaging personality, reliability, and so on). Don’t overdo it—six dimensions is a good number. The traits you choose should be as independent as possible from each other, and you should feel that you can assess them reliably by asking a few factual questions. Next, make a list of those questions for each trait and think about how you will score it, say on a 1–5 scale. You should have an idea of what you will call “very weak” or “very strong.”

These preparations should take you half an hour or so, a small investment that can make a significant difference in the quality of the people you hire. To avoid halo effects, you must collect the information on one trait at a time, scoring each before you move on to the next one. Do not skip around. To evaluate each candidate, add up the six scores. Because you are in charge of the final decision, you should not do a “close your eyes.” Firmly resolve that you will hire the candidate whose final score is the highest, even

if there is another one whom you like better—try to resist your wish to invent broken legs to change the ranking. A vast amount of research offers a promise: you are much more likely to find the best candidate if you use this procedure than if you do what people normally do in such situations, which is to go into the interview unprepared and to make choices by an overall intuitive judgment such as “I looked into his eyes and liked what I saw.”

SPEAKING OF JUDGES VS. FORMULAS

“Whenever we can replace human judgment by a formula, we should at least consider it.”

“He thinks his judgments are complex and subtle, but a simple combination of scores could probably do better.”

“Let’s decide in advance what weight to give to the data we have on the candidates’ past performance. Otherwise we will give too much weight to our impression from the interviews.”

experiments rather than by the study of real people doing things that matter. They are deeply skeptical about the value of using rigid algorithms to replace human judgment, and Paul Meehl is not among their heroes. Gary Klein has eloquently articulated this position over many years.

This is hardly the basis for a beautiful friendship, but there is more to the story. I had never believed that intuition is always misguided. I had also been a fan of Klein's studies of expertise in firefighters since I first saw a draft of a paper he wrote in the 1970s, and was impressed by his book *Sources of Power*, much of which analyzes how experienced professionals develop intuitive skills. I invited him to join in an effort to map the boundary that separates the marvels of intuition from its flaws. He was intrigued by the idea and we went ahead with the project—with no certainty that it would succeed. We set out to answer a specific question: When can you trust an experienced professional who claims to have an intuition? It was obvious that Klein would be more disposed to be trusting, and I would be more skeptical. But could we agree on principles for answering the general question?

Over seven or eight years we had many discussions, resolved many disagreements, almost blew up more than once, wrote many drafts, became friends, and eventually published a joint article with a title that tells the story: "Conditions for Intuitive Expertise: A Failure to Disagree." Indeed, we did not encounter real issues on which we disagreed—but we did not really agree.

MARVELS AND FLAWS

Malcolm Gladwell's bestseller *Blink* appeared while Klein and I were working on the project, and it was reassuring to find ourselves in agreement about it. Gladwell's book opens with the memorable story of art experts faced with an object that is described as a magnificent example of a kouros, a sculpture of a striding boy. Several of the experts had strong visceral reactions: they felt in their gut that the statue was a fake but were not able to articulate what it was about it that made them uneasy. Everyone who read the book—millions did—remembers that story as a triumph of intuition. The experts agreed that they knew the sculpture was a fake without knowing how they knew—the very definition of intuition. The story appears to imply that a systematic search for the cue that guided the experts would have failed, but Klein and I both rejected that conclusion. From our point of view, such an inquiry was needed, and if it had been conducted properly (which Klein knows how to do), it would probably have succeeded.

EXPERT INTUITION: WHEN CAN WE TRUST IT?

Professional controversies bring out the worst in academics. Scientific journals occasionally publish exchanges, often beginning with someone's critique of another's research, followed by a reply and a rejoinder. I have always thought that these exchanges are a waste of time. Especially when the original critique is sharply worded, the reply and the rejoinder are often exercises in what I have called sarcasm for beginners and advanced sarcasm. The replies rarely concede anything to a biting critique, and it is almost unheard of for a rejoinder to admit that the original critique was misguided or erroneous in any way. On a few occasions I have responded to criticisms that I thought were grossly misleading, because a failure to respond can be interpreted as conceding error, but I have never found the hostile exchanges instructive. In search of another way to deal with disagreements, I have engaged in a few "adversarial collaborations," in which scholars who disagree on the science agree to write a jointly authored paper on their differences, and sometimes conduct research together. In especially tense situations, the research is moderated by an arbiter.

My most satisfying and productive adversarial collaboration was with Gary Klein, the intellectual leader of an association of scholars and practitioners who do not like the kind of work I do. They call themselves students of Naturalistic Decision Making, or NDM, and mostly work in organizations where they often study how experts work. The NDMers adamantly reject the focus on biases in the heuristics and biases approach. They criticize this model as overly concerned with failures and driven by artificial

Although many readers of the kourou example were surely drawn to an almost magical view of expert intuition, Gladwell himself does not hold that position. In a later chapter he describes a massive failure of intuition: Americans elected President Harding, whose only qualification for the position was that he perfectly looked the part. Square jawed and tall, he was the perfect image of a strong and decisive leader. People voted for someone who looked strong and decisive without any other reason to believe that he was. An intuitive prediction of how Harding would perform as president arose from substituting one question for another. A reader of this book should expect such an intuition to be held with confidence.

INTUITION AS RECOGNITION

The early experiences that shaped Klein's views of intuition were starkly different from mine. My thinking was formed by observing the illusion of validity in myself and by reading Paul Meehl's demonstrations of the inferiority of clinical prediction. In contrast, Klein's views were shaped by his early studies of fireground commanders (the leaders of firefighting teams). He followed them as they fought fires and later interviewed the leader about his thoughts as he made decisions. As Klein described it in our joint article, he and his collaborators

investigated how the commanders could make good decisions without comparing options. The initial hypothesis was that commanders would restrict their analysis to only a pair of options, but that hypothesis proved to be incorrect. In fact, the commanders usually generated only a single option, and that was all they needed. They could draw on the repertoire of patterns that they had compiled during more than a decade of both real and virtual experience to identify a plausible option, which they considered first. They evaluated this option by mentally simulating it to see if it would work in the situation they were facing. . . . If the course of action they were considering seemed appropriate, they would implement it. If it had shortcomings, they would modify it. If they could not easily modify it, they would turn to the next most plausible option and run through the same procedure until an acceptable course of action was found.

Klein elaborated this description into a theory of decision making that he called the recognition-primed decision (RPD) model, which applies to firefighters but also describes expertise in other domains, including chess. The

process involves both System 1 and System 2. In the first phase, a tentative plan comes to mind by an automatic function of associative memory—System 1. The next phase is a deliberate process in which the plan is mentally simulated to check if it will work—an operation of System 2. The model of intuitive decision making as pattern recognition develops ideas presented some time ago by Herbert Simon, perhaps the only scholar who is recognized and admired as a hero and founding figure by all the competing clans and tribes in the study of decision making. I quoted Herbert Simon's definition of intuition in the introduction, but it will make more sense when I repeat it now: "The situation has provided a cue; this cue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition."

This strong statement reduces the apparent magic of intuition to the everyday experience of memory. We marvel at the story of the firefighter who has a sudden urge to escape a burning house just before it collapses, because the firefighter knows the danger intuitively, "without knowing how he knows." However, we also do not know how we immediately know that a person we see as we enter a room is our friend Peter. The moral of Simon's remark is that the mystery of knowing without knowing is not a distinctive feature of intuition; it is the norm of mental life.

ACQUIRING SKILL

How does the information that supports intuition get "stored in memory"? Certain types of intuitions are acquired very quickly. We have inherited from our ancestors a great facility to learn when to be afraid. Indeed, one experience is often sufficient to establish a long-term aversion and fear. Many of us have the visceral memory of a single dubious dish that still leaves us vaguely reluctant to return to a restaurant. All of us tense up when we approach a spot in which an unpleasant event occurred, even when there is no reason to expect it to happen again. For me, one such place is the ramp leading to the San Francisco airport, where years ago a driver in the throes of road rage followed me from the freeway, rolled down his window, and hurled obscenities at me. I never knew what caused his hatred, but I remember his voice whenever I reach that point on my way to the airport.

My memory of the airport incident is conscious and it fully explains the emotion that comes with it. On many occasions, however, you may feel uneasy in a particular place or when someone uses a particular turn of phrase without having a conscious memory of the triggering event. In hindsight,

you will label that unease an intuition if it is followed by a bad experience. This mode of emotional learning is closely related to what happened in Pavlov's famous conditioning experiments, in which the dogs learned to recognize the sound of the bell as a signal that food was coming. What Pavlov's dogs learned can be described as a learned hope. Learned fears are even more easily acquired.

Fear can also be learned—quite easily, in fact—by words rather than by experience. The fireman who had the “sixth sense” of danger had certainly had many occasions to discuss and think about types of fires he was not involved in, and to rehearse in his mind what the cues might be and how he should react. As I remember from experience, a young platoon commander with no experience of combat will tense up while leading troops through a narrowing ravine, because he was taught to identify the terrain as favoring an ambush. Little repetition is needed for learning.

Emotional learning may be quick, but what we consider as “expertise” usually takes a long time to develop. The acquisition of expertise in complex tasks such as high-level chess, professional basketball, or firefighting is intricate and slow because expertise in a domain is not a single skill but rather a large collection of miniskills. Chess is a good example. An expert player can understand a complex position at a glance, but it takes years to develop that level of ability. Studies of chess masters have shown that at least 10,000 hours of dedicated practice (about 6 years of playing chess 5 hours a day) are required to attain the highest levels of performance. During those hours of intense concentration, a serious chess player becomes familiar with thousands of configurations, each consisting of an arrangement of related pieces that can threaten or defend each other.

Learning high-level chess can be compared to learning to read. A first grader works hard at recognizing individual letters and assembling them into syllables and words, but a good adult reader perceives entire clauses. An expert reader has also acquired the ability to assemble familiar elements in a new pattern and can quickly “recognize” and correctly pronounce a word that she has never seen before. In chess, recurrent patterns of interacting pieces play the role of letters, and a chess position is a long word or a sentence.

A skilled reader who sees it for the first time will be able to read the opening stanza of Lewis Carroll's “Jabberwocky” with perfect rhythm and intonation, as well as pleasure:

“Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:

All mimsy were the borogoves,
And the mome raths outgrabe.

Acquiring expertise in chess is harder and slower than learning to read because there are many more letters in the “alphabet” of chess and because the “words” consist of many letters. After thousands of hours of practice, however, chess masters are able to read a chess situation at a glance. The few moves that come to their mind are almost always strong and sometimes creative. They can deal with a “word” they have never encountered, and they can find a new way to interpret a familiar one.

THE ENVIRONMENT OF SKILL

Klein and I quickly found that we agreed both on the nature of intuitive skill and on how it is acquired. We still needed to agree on our key question: When can you trust a self-confident professional who claims to have an intuition?

We eventually concluded that our disagreement was due in part to the fact that we had different experts in mind. Klein had spent much time with fireground commanders, clinical nurses, and other professionals who have real expertise. I had spent more time thinking about clinicians, stock pickers, and political scientists trying to make unsupportable long-term forecasts. Not surprisingly, his default attitude was trust and respect; mine was skepticism. He was more willing to trust experts who claim an intuition because, as he told me, true experts know the limits of their knowledge. I argued that there are many pseudo-experts who have no idea that they do not know what they are doing (the illusion of validity), and that as a general proposition subjective confidence is commonly too high and often uninformative.

Earlier I traced people's confidence in a belief to two related impressions: cognitive ease and coherence. We are confident when the story we tell ourselves comes easily to mind, with no contradiction and no competing scenario. But ease and coherence do not guarantee that a belief held with confidence is true. The associative machine is set to suppress doubt and to evoke ideas and information that are compatible with the currently dominant story. A mind that follows WYSIATI will achieve high confidence much too easily by ignoring what it does not know. It is therefore not surprising that many of us are prone to have high confidence in unfounded intuitions. Klein and I eventually agreed on an important principle: the

confidence that people have in their intuitions is not a reliable guide to their validity. In other words, do not trust anyone—including yourself—to tell you how much you should trust their judgment.

If subjective confidence is not to be trusted, how can we evaluate the probable validity of an intuitive judgment? When do judgments reflect true expertise? When do they display an illusion of validity? The answer comes from the two basic conditions for acquiring a skill:

- an environment that is sufficiently regular to be predictable
- an opportunity to learn these regularities through prolonged practice

When both these conditions are satisfied, intuitions are likely to be skilled. Chess is an extreme example of a regular environment, but bridge and poker also provide robust statistical regularities that can support skill. Physicians, nurses, athletes, and firefighters also face complex but fundamentally orderly situations. The accurate intuitions that Gary Klein has described are due to highly valid cues that the expert's System 1 has learned to use, even if System 2 has not learned to name them. In contrast, stock pickers and political scientists who make long-term forecasts operate in a zero-validity environment. Their failures reflect the basic unpredictability of the events that they try to forecast.

Some environments are worse than irregular. Robin Hogarth described “wicked” environments, in which professionals are likely to learn the wrong lessons from experience. He borrows from Lewis Thomas the example of a physician in the early twentieth century who often had intuitions about patients who were about to develop typhoid. Unfortunately, he tested his hunch by palpating the patient's tongue, without washing his hands between patients. When patient after patient became ill, the physician developed a sense of clinical infallibility. His predictions were accurate—but not because he was exercising professional intuition!

Meehl's clinicians were not inept and their failure was not due to lack of talent. They performed poorly because they were assigned tasks that did not have a simple solution. The clinicians' predicament was less extreme than the zero-validity environment of long-term political forecasting, but they operated in low-validity situations that did not allow high accuracy. We know this to be the case because the best statistical algorithms, although more accurate than human judges, were never very accurate. Indeed, the

studies by Meehl and his followers never produced a “smoking gun” demonstration, a case in which clinicians completely missed a highly valid cue that the algorithm detected. An extreme failure of this kind is unlikely because human learning is normally efficient. If a strong predictive cue exists, human observers will find it, given a decent opportunity to do so. Statistical algorithms greatly outdo humans in noisy environments for two reasons: they are more likely than human judges to detect weakly valid cues and much more likely to maintain a modest level of accuracy by using such cues consistently.

It is wrong to blame anyone for failing to forecast accurately in an unpredictable world. However, it seems fair to blame professionals for believing they can succeed in an impossible task. Claims for correct intuitions in an unpredictable situation are self-delusional at best, sometimes worse. In the absence of valid cues, intuitive “hits” are due either to luck or to lies. If you find this conclusion surprising, you still have a lingering belief that intuition is magic. Remember this rule: intuition cannot be trusted in the absence of stable regularities in the environment.

FEEDBACK AND PRACTICE

Some regularities in the environment are easier to discover and apply than others. Think of how you developed your style of using the brakes on your car. As you were mastering the skill of taking curves, you gradually learned when to let go of the accelerator and when and how hard to use the brakes. Curves differ, and the variability you experienced while learning ensures that you are now ready to brake at the right time and strength for any curve you encounter. The conditions for learning this skill are ideal, because you receive immediate and unambiguous feedback every time you go around a bend: the mild reward of a comfortable turn or the mild punishment of some difficulty in handling the car if you brake either too hard or not quite hard enough. The situations that face a harbor pilot maneuvering large ships are no less regular, but skill is much more difficult to acquire by sheer experience because of the long delay between actions and their noticeable outcomes. Whether professionals have a chance to develop intuitive expertise depends essentially on the quality and speed of feedback, as well as on sufficient opportunity to practice.

Expertise is not a single skill; it is a collection of skills, and the same professional may be highly expert in some of the tasks in her domain while remaining a novice in others. By the time chess players become experts,

they have “seen everything” (or almost everything), but chess is an exception in this regard. Surgeons can be much more proficient in some operations than in others. Furthermore, some aspects of any professional’s tasks are much easier to learn than others. Psychotherapists have many opportunities to observe the immediate reactions of patients to what they say. The feedback enables them to develop the intuitive skill to find the words and the tone that will calm anger, forge confidence, or focus the patient’s attention. On the other hand, therapists do not have a chance to identify which general treatment approach is most suitable for different patients. The feedback they receive from their patients’ long-term outcomes is sparse, delayed, or (usually) nonexistent, and in any case too ambiguous to support learning from experience.

Among medical specialties, anesthesiologists benefit from good feedback, because the effects of their actions are likely to be quickly evident. In contrast, radiologists obtain little information about the accuracy of the diagnoses they make and about the pathologies they fail to detect. Anesthesiologists are therefore in a better position to develop useful intuitive skills. If an anesthesiologist says, “I have a feeling something is wrong,” everyone in the operating room should be prepared for an emergency.

Here again, as in the case of subjective confidence, the experts may not know the limits of their expertise. An experienced psychotherapist knows that she is skilled in working out what is going on in her patient’s mind and that she has good intuitions about what the patient will say next. It is tempting for her to conclude that she can also anticipate how well the patient will do next year, but this conclusion is not equally justified. Short-term anticipation and long-term forecasting are different tasks, and the therapist has had adequate opportunity to learn one but not the other. Similarly, a financial expert may have skills in many aspects of his trade but not in picking stocks, and an expert in the Middle East knows many things but not the future. The clinical psychologist, the stock picker, and the pundit do have intuitive skills in some of their tasks, but they have not learned to identify the situations and the tasks in which intuition will betray them. The unrecognized limits of professional skill help explain why experts are often overconfident.

EVALUATING VALIDITY

At the end of our journey, Gary Klein and I agreed on a general answer to our initial question: When can you trust an experienced professional who

claims to have an intuition? Our conclusion was that for the most part it is possible to distinguish intuitions that are likely to be valid from those that are likely to be bogus. As in the judgment of whether a work of art is genuine or a fake, you will usually do better by focusing on its provenance than by looking at the piece itself. If the environment is sufficiently regular and if the judge has had a chance to learn its regularities, the associative machinery will recognize situations and generate quick and accurate predictions and decisions. You can trust someone’s intuitions if these conditions are met.

Unfortunately, associative memory also generates subjectively compelling intuitions that are false. Anyone who has watched the chess progress of a talented youngster knows well that skill does not become perfect all at once, and that on the way to near perfection some mistakes are made with great confidence. When evaluating expert intuition you should always consider whether there was an adequate opportunity to learn the cues, even in a regular environment.

In a less regular, or low-validity, environment, the heuristics of judgment are invoked. System 1 is often able to produce quick answers to difficult questions by substitution, creating coherence where there is none. The question that is answered is not the one that was intended, but the answer is produced quickly and may be sufficiently plausible to pass the lax and lenient review of System 2. You may want to forecast the commercial future of a company, for example, and believe that this is what you are judging, while in fact your evaluation is dominated by your impressions of the energy and competence of its current executives. Because substitution occurs automatically, you often do not know the origin of a judgment that you (your System 2) endorse and adopt. If it is the only one that comes to mind, it may be subjectively undistinguishable from valid judgments that you make with expert confidence. This is why subjective confidence is not a good diagnostic of accuracy: judgments that answer the wrong question can also be made with high confidence.

You may be asking, Why didn’t Gary Klein and I come up immediately with the idea of evaluating an expert’s intuition by assessing the regularity of the environment and the expert’s learning history—mostly setting aside the expert’s confidence? And what did we think the answer could be? These are good questions because the contours of the solution were apparent from the beginning. We knew at the outset that fireground commanders and pediatric nurses would end up on one side of the boundary of valid intuitions and that the specialties studied by Meehl would be on the other, along with stock pickers and pundits.

It is difficult to reconstruct what it was that took us years, long hours of discussion, endless exchanges of drafts and hundreds of e-mails negotiating over words, and more than once almost giving up. But this is what always happens when a project ends reasonably well: once you understand the main conclusion, it seems it was always obvious.

As the title of our article suggests, Klein and I disagreed less than we had expected and accepted joint solutions of almost all the substantive issues that were raised. However, we also found that our early differences were more than an intellectual disagreement. We had different attitudes, emotions, and tastes, and those changed remarkably little over the years. This is most obvious in the facts that we find amusing and interesting. Klein still winces when the word *bias* is mentioned, and he still enjoys stories in which algorithms or formal procedures lead to obviously absurd decisions. I tend to view the occasional failures of algorithms as opportunities to improve them. On the other hand, I find more pleasure than Klein does in the commonplace of arrogant experts who claim intuitive powers in zero-validity situations. In the long run, however, finding as much intellectual agreement as we did is surely more important than the persistent emotional differences that remained.

SPEAKING OF EXPERT INTUITION

"How much expertise does she have in this particular task? How much practice has she had?"

"Does he really believe that the environment of start-ups is sufficiently regular to justify an intuition that goes against the base rates?"

"She is very confident in her decision, but subjective confidence is a poor index of the accuracy of a judgment."

"Did he really have an opportunity to learn? How quick and how clear was the feedback he received on his judgments?"